

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

B5

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : A61K 39/00, C12N 15/12, C12P 21/02	A1	(11) International Publication Number: WO 95/35118 (43) International Publication Date: 28 December 1995 (28.12.95)
(21) International Application Number: PCT/US95/07295 (22) International Filing Date: 6 June 1995 (06.06.95) (30) Priority Data: 264,101 20 June 1994 (20.06.94) US (60) Parent Application or Grant (63) Related by Continuation US 264,101 (CIP) Filed on 20 June 1994 (20.06.94) (71) Applicant (for all designated States except US): MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). (72) Inventors; and (75) Inventors/Applicants (for US only): ALVES, Kenneth [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). GUPTA, Sunil, K. [IN/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US). HOLLIS, Gregory, Franklin [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).	(74) Common Representative: MERCK & CO., INC.; Patent Dept., 126 East Lincoln Avenue, Rahway, NJ 07065 (US). (81) Designated States: AM, AU, BB, BG, BR, BY, CA, CN, CZ, EE, FI, GE, HU, IS, JP, KG, KR, KZ, LK, LR, LT, LV, MD, MG, MN, MX, NO, NZ, PL, RO, RU, SG, SI, SK, TJ, TM, TT, UA, US, UZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG). Published <i>With international search report.</i>	
(54) Title: CONTRACEPTIVE VACCINE (57) Abstract <p>The instant invention is drawn to a sperm surface protein in substantially pure form selected from a human PH30 beta chain protein and a mouse PH30 beta chain proteins. Such proteins are useful as contraceptive vaccines in humans and mice respectively, and for identifying small molecules that will disrupt sperm-egg interaction and fertilization.</p>		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

- 1 -

TITLE OF THE INVENTION
CONTRACEPTIVE VACCINE

FIELD OF THE INVENTION

5 The present invention provides sperm surface proteins and DNA sequences encoding the proteins which are useful in the prevention of fertilization. More particularly, the cloning and characterization of the mouse and human PH3C beta chain genes, as well as their use as contraceptive vaccines, are described.

10

BACKGROUND OF THE INVENTION

 Four methods of family planning are currently available in the U.S., sterilization, abstinence, abortion and contraception. Of these four birth control methods, contraception is the most widely utilized.

15 Despite the substantial U.S. and global demand for contraception, the presently available methodologies fall short of market needs. Oral contraceptives and barrier methods dominate today's contraceptive market but have significant shortcomings. Oral contraceptives, though efficacious, are documented to be associated with significant side effects

20 including increased risks of cardiovascular disease and breast cancer and are not recommended for women over the age of 35. Barrier methods, while safe, have failure rates approaching 20%. There is a clear need for increased availability of and improvements in contraceptives that offer superior safety, efficacy, convenience, acceptability and are

25 affordable to women and men worldwide. Identification of novel approaches for controlling fertility is therefore necessary.

 Immunization of male and female animals with extracts of whole sperm is known to cause infertility. [Tung, K., et al., *J. Reproductive Immunol.*, 1; 145-158 (1979); Menge, A., et al., *Biol. of*

30 *Reproduction*, 20, 931-937 (1979)]. Moreover, men and women who spontaneously produce antisperm antibodies are infertile, but otherwise healthy. [Bronson, R., et al., *Fert. and Sterile*, 42, 171-183 (1984)].

- 2 -

Although the critical sperm antigens are unknown, these observations have led to the proposal that sperm proteins might be useful in the development of a contraceptives vaccine.

In mammalian species, sperm proteins are believed to have a role in sperm adhesion to the zona pellucida of the egg. The PH30 protein is known to be involved in sperm egg binding and antibodies that bind to PH30 inhibit this interaction. PH30 is an integral membrane protein present on posterior head of sperm which mediates sperm-oocyte fusion. The PH30 protein consists of two immunologically distinct alpha and beta subunits. Both subunits are made as larger precursors and then finally processed in epididymis where sperm become fertilization competent. [Primakoff, P., et al., *J. Cell Biology*, 104, 141-149 (1987); Blobel, C.P., et al., *J. Cell Biology*, 111, 69-78 (1990)]. Monoclonal antibodies that recognize PH30 inhibit sperm-oocyte fusion *in vitro*, indicating its importance in fertilization [Primakoff, P., et al., *J. Cell Biology*, 104, 141-149 (1987)].

Guinea pig PH30 alpha and beta chains have been cloned by Blobel et al. Mature PH30 alpha chain consists of 289 amino acids and encodes a transmembrane domain as well as an integral fusion peptide (82-102) that is similar to a potential fusion peptide of E2 glycoprotein of rubella virus. Guinea Pig PH30 beta chain has an open reading frame of 353 amino acids and also encodes a transmembrane domain. [Blobel C.P., et al., *Nature*, 356, 248-251 (1992)]. The predicted amino acid sequence of the PH30 beta chain protein contains significant homology to a class of proteins called disintegrins found in snake venom. These proteins are known to bind to a family of proteins called integrins and prevent their normal functioning in cell adhesion (a well studied example is platelet aggregation). The N-terminal ninety amino acids integrin binding disintegrin domain of PH30 beta has been postulated to mediate the binding of PH30 to its putative integrin receptor on oocytes. The cloning and sequence determination of the mouse and human PH30 beta chain genes would permit novel approaches to the control of sperm

- 3 -

egg binding and fusions. These approaches include, but are not limited to, eliciting an immune response directed at all or part of the PH30 beta chain protein and using the PH30 beta chain protein as part of a screen to identify small molecules that alter sperm egg interactions.

5 Mammalian fertilization is, in most cases, species specific. Thus, the identification and isolation of sperm surface proteins essential for fertilization in species other than guinea pig would be useful for providing effective long lasting contraception in those species. Thus far, the lack of biochemical identification, isolation and cloning of
10 candidate adhesion proteins of sperm has hindered scientists in developing effective contraceptives for humans as well as other mammalian species.

SUMMARY OF THE INVENTION

15 The instant invention relates to a sperm protein in substantially pure form selected from a human PH30 beta chain protein, a mouse PH30 beta chain protein or an amino acid sequence substantially homologous to either the human or mouse PH30 beta chain protein.

20 In one embodiment of the invention is the sperm protein having an integrin binding sequence which is not TDE.

In one class is the sperm protein wherein the integrin binding sequence is selected from FEE or QDE.

In a subclass is the sperm protein which is the human PH30
25 beta chain protein.

Illustrative of this subclass is the sperm protein having an integrin binding sequence that is FEE.

Further illustrating the invention is a DNA sequence which encodes the sperm protein or a portion of the sperm protein sufficient
30 to constitute at least one epitope.

An illustration is the DNA sequence wherein the epitope is on the native protein.

- 4 -

Exemplifying the invention is the DNA sequence which encodes all or a portion of human PH30 beta chain protein.

An example of the invention is the DNA sequence, wherein the DNA encoding all or a portion of the human PH30 beta protein is characterized by the ability to hybridize, under standard conditions, to the DNA sequence shown in SEQ ID NO: 1.

More particularly illustrating the invention is a contraceptive composition comprising a therapeutically effective amount of the protein, or a polypeptide having the substantially same amino acid sequence as a segment of the protein provided that the polypeptide is sufficient to constitute at least one epitope, and a pharmaceutically acceptable carrier.

Another illustration is the contraceptive composition wherein the epitope is on the native protein.

Further exemplifying the invention is the contraceptive composition, wherein the protein is the human PH30 beta chain protein.

More specifically illustrating the invention is the contraceptive composition, wherein the protein is produced by expressing the gene encoding an immunogenic epitope of the sperm protein in a recombinant DNA expression vector.

Specifically exemplifying the invention is a vector comprising an inserted DNA sequence encoding for the protein.

A further illustration of the invention is the vector, wherein the inserted DNA sequence is characterized by the ability to hybridize, under standard conditions, to a DNA sequence selected from the DNA sequences of SEQ ID NO: 1 or SEQ ID NO: 3.

Another example of the invention is a host that is compatible with and contains the vector.

More specifically exemplifying the invention is a method of producing a human or mouse PH30 beta chain sperm protein, comprising the steps of culturing cells containing PH30 beta chain DNA and recovering the sperm protein from the cell culture.

- 5 -

A further example is the method wherein the DNA encoding all or a portion of the PH30 beta chain protein is characterized by the ability to hybridize, under standard conditions, to a DNA sequence selected from the DNA sequences of SEQ ID NO: 1 or SEQ ID NO: 3.

A more specific illustration is a method of contraception in a human or mouse subject in need thereof, comprising administering to the subject an amount of the sperm protein which is effective for the stimulation of antibodies which bind to the sperm protein in vivo, thereby preventing or substantially reducing the rate of sperm-egg fusion.

Further illustrating the invention is the method wherein the sperm protein has an integrin binding sequence which is not TDE.

Another illustration is the PH30 beta chain protein made by the process described.

Another example is a DNA sequence as shown in Seq. ID No. 1 encoding human PH30 beta chain protein.

Still further illustrating the invention is a purified and isolated DNA sequence consisting essentially of a DNA sequence encoding a polypeptide having an amino acid sequence sufficiently duplicative of that of human or mouse PH30 beta to allow the possession of the biological property of initiating sperm-egg binding or promoting sperm-egg fusion. This biological activity can be determined using the in vitro sperm-oocyte binding/fusion assays [Primakoff, P., et al., *J. Cell. Biol.*, 104: 141-149 (1987)].

More particularly exemplifying the invention is the DNA sequence wherein the amino acid sequence contains an integrin binding sequence which is not TDE.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a diagram representing the human PH30 beta cDNA gene sequence encoding the human PH-30 beta protein, and the

- 6 -

deduced amino acid sequence of the human PH-30 beta protein present in three letter code. The sequence disclosure of Figure 1 is represented as SEQ ID NO: 1 and 2.

Figure 2 is a diagram representing the mouse PH30 beta
5 cDNA gene sequence, and the deduced amino acid sequence of the mouse PH-30 beta protein present in three letter code. The sequence disclosure of Figure 2 is represented as SEQ ID NO: 3 and 4.

Figure 3 is a restriction MAP of the human PH30 beta cDNA sequence.

10 Figure 4 is a restriction MAP of the mouse PH30 beta cDNA sequence.

DETAILED DESCRIPTION OF THE INVENTION

The subject invention relates to sperm surface proteins which
15 are essential for fertilization, or portions thereof, and their use in contraceptive methods. A sperm surface protein is essential for fertilization if, for example, a monoclonal antibody to the protein or a polyclonal antibody raised against the purified protein, when bound to sperm, inhibits *in vitro* or *in vivo* fertilization or any step of *in vitro*
20 fertilization. The process of fertilization is defined as the binding or fusion of two gametes (sperm and egg) followed by the fusion of their nuclei to form the genome of a new organism. The surface protein can be located in the plasma membrane of sperm and/or the inner acrosomal membrane. It can be a protein or glycoprotein. The isolated surface protein used for
25 immunization can comprise the entire surface protein or some portion of the protein (external to the cell) which is immunogenic. Two such sperm surface proteins are the mouse and human PH30 beta chain sperm surface proteins. The PH30 beta genes encode proteins which are present on the surface of sperm cells and are essential for fertilization.

30 As used herein, a protein or peptide is "substantially pure" when that protein or peptide has been purified to the extent that it is essentially free of other molecules with which it is associated in nature.

- 7 -

The term "substantially pure" is used relative to proteins or peptides with which the peptides of the instant invention are associated in nature, and are not intended to exclude compositions in which the peptide of the invention is admixed with nonproteinous pharmaceutical carriers or vehicles.

5 As used herein, an amino acid sequence substantially homologous to a referent PH-30 beta protein will have at least 70% sequence homology, preferably 80%, and most preferably 90% sequence homology with the amino acid sequence of a referent PH-30 beta protein or a peptide thereof. For example, an amino acid sequence is substantially
10 homologous to mouse PH-30 beta protein if, when aligned with mouse PH-30 beta protein, at least 70% of its amino acid residues are the same. In addition, it is preferable that the substantially homologous amino acid sequence contains the integrin binding sequence.

 As used herein, a DNA sequence substantially homologous to a
15 referent PH-30 beta protein will have at least 70%, preferably 80%, and most preferably 90% sequence homology with the DNA sequence of a referent PH-30 beta. Moreover, a DNA sequence substantially homologous to a referent PH-30 beta protein is characterized by the ability to hybridize to the DNA sequence of a referent PH30 beta under standard conditions.
20 Standard hybridization conditions are described in Maniatis, T., et al. (1989) Molecular Cloning, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

 An "expression vector" or "vector," as used herein, refers to a plasmid, bacteriophage, virus, or other molecule into which a gene of
25 interest may be cloned, such that the appropriate signals for expression of that gene are present on that vector.

 The term "epitope," as used herein, refers to the minimum amount of PH30 beta sequence capable of producing an efficacious, i.e., contraceptive, immune response.

30 The term "therapeutically effective amount," as used herein, means that amount of a drug or pharmaceutical agent that will elicit the

- 8 -

biological or medical response that is being sought by a researcher or clinician.

Production and Purification of Immunogen

5 A preferred method for producing sperm surface proteins for use as a contraceptive immunogen is by recombinant DNA technology. To produce the protein using this technology it is necessary to isolate and clone DNA encoding the protein, or an immunogenic portion thereof. Those skilled in the art are familiar with a variety of
10 approaches which can be used in an effort to clone a gene of interest. However, having nothing more than the isolated protein of interest, success in such an effort cannot be predicted with a reasonable degree of certainty.

 In the Examples which follow, Applicants describe the
15 cloning and characterization of the mouse and human PH30 beta chain genes. The mouse and human PH30 beta chain genes were isolated using a cDNA encoding the guinea pig PH30 beta chain gene. The instant invention provides specific sequence information to permit targeted intervention in controlling fertility through anti PH30 directed
20 immune responses inhibition of sperm-egg binding and triggering of post binding signaling and effective events. These sequences permit the generation of reagents for the isolation of oocyte proteins involved in sperm-egg interaction.

 The information presented in the Examples enable one
25 skilled in the art to isolate and clone the mouse or human PH30 beta chain gene. For example, a cDNA library is prepared from testis or spermatogenic cells isolated from the mammal of interest (e.g., mouse, human). Such a cDNA library is then screened using, for example, labeled guinea pig PH30 DNA probes. DNA encoding all or a portion
30 of human or mouse PH30 is characterized by the ability to hybridize to such a probe sequence under hybridization conditions such as those described in Example 1. Methods of labeling and screening by

- 9 -

hybridization are well known in the art. Positive clones are analyzed, and a full length cDNA is constructed by conventional methods.

The cloned gene, or portions thereof which encode an immunogenic region of the PH30 protein, can be expressed by inserting the coding region into an expression vector to produce an expression construct. Many such expression vectors are known to those skilled in the art. These vectors contain a promoter for the gene of interest as well as additional transcriptional and translational signals. Expression vectors for both eukaryotic host cells and prokaryotic host cells are widely available. The DNA expression construct is used to transform an appropriate host cell.

Eukaryotic, in particular mammalian, host cells are often utilized for the expression of eukaryotic proteins. It has been found, for example, that eukaryotic proteins may exhibit folding problems when expressed in prokaryotic cells. In addition, production of authentic, biologically active eukaryotic proteins from cloned DNA sometimes requires post-translational modification such as disulfide bond formation, glycosylation, phosphorylation or specific proteolytic cleavage processes that are not performed in bacterial cells. This is especially true with membrane proteins. The sperm surface protein is produced using the transcriptional and translational components of the host cell. After an appropriate growth and expression period, the host cell culture is lysed and the sperm surface protein is purified from the lysate. Lysis buffers typically include non-ionic detergent, protease inhibitors, etc.

From the solubilized cell extract, the sperm surface protein can be purified and isolated by physical and biochemical methods such as ultracentrifugation, column chromatography, high performance liquid chromatography, electrophoresis, etc. Alternatively, the sperm surface protein can be isolated by affinity chromatography using monoclonal or polyclonal antibodies [see Primakoff et al., *Biol. of*

- 10 -

Reprod. 38, 921-934 (1988)]. Such methods for purifying proteins are well known to those skilled in the art.

As mentioned above, antigenic portions or epitopes of the sperm surface protein are useful as immunogen, in addition to the full length protein. Antigenic fragments can be produced, for example, by proteolytic digestion of the full length protein, followed by isolation of the desired fragment. Alternatively, chemical synthesis can be used to generate the desired fragment starting with monomer amino acid residues.

With respect to the PH30 protein, certain antigenic domains are preferred candidates for use in a contraceptive vaccine. As is discussed in greater detail in the Exemplification section which follows, the PH30 β subunit contains a domain which is highly conserved when compared to a class of proteins known as disintegrins. A peptide (or portion thereof) which is identical or substantially identical to this domain is preferred for use in the contraceptive methods of this invention. Substantially identical, as used in the preceding sentence, means that at least 70% of the amino acid sequence of the peptide is identical to the corresponding portion of the PH30 β disintegrin domain.

Disintegrins are found in snake venom, for example, and are known to bind to a class of platelet surface proteins known as integrins. The binding of disintegrins to integrins has been shown to inhibit blood clotting. By analogy, peptides corresponding to the PH30 β disintegrin domain are predicted to be active in sperm-egg binding and fusion.

Contraceptive Vaccine

Once the sperm surface protein has been produced and purified, a vaccine can be produced by combining the sperm surface protein or portion thereof with a suitable carrier for administration to a subject for immunization. For successful vaccine development it is

- 11 -

necessary that the immunogen exhibit tissue specificity, that is, it is expressed on the target tissue only and must be essential for the process of reproduction. It is known that the PH30 protein, which is expressed only on sperm, is involved in sperm egg binding and antibodies that
5 bind to PH30 inhibit that interaction.

The cloning and characterization of human PH30 beta permits novel approaches for using PH30 as a target to control human fertility. PH30 beta protein or peptides can be used directly as an antigen to elicit an immune response directed to the whole or a relevant
10 part of the PH30 beta chain protein. Testing of these approaches requires availability of sufficient quantities of PH30 beta protein. The cloning and sequencing of the mouse and human PH30 beta chain provides information necessary to recombinantly express all or part of the PH30 beta protein. These expressed proteins are used with or
15 without adjuvant to immunize women or female mice. The elicited humoral immune responses are monitored by assays that use PH30 beta as antigen. Secreted antibodies in the female reproductive system will bind to the sperm head and disrupt fertilization. The availability of the recombinant mouse PH30 beta protein permits establishment of an
20 animal model system for testing efficacy, reversibility and safety of specific methods of controlling fertility based on PH30.

A vaccine can contain one or more sperm surface proteins. Sperm surface proteins of the present invention can be combined with adjuvants which contain non-specific stimulators of the immune system.
25 Proper use of adjuvants can induce a strong antibody response to foreign antigens (i.e., sperm surface proteins). The action of adjuvants is not fully understood, but most adjuvants incorporate two components. One is a substance designed to form a deposit which protects the antigen from catabolism. Two methods of forming a deposit are to use mineral
30 oils or aluminum hydroxide precipitates. With mineral oils, such as Freund's adjuvant, the immunogen is prepared in a water-in-oil

- 12 -

emulsion. For aluminum hydroxide, the immunogen is either adsorbed to preformed precipitants or is trapped during precipitation.

The second component required for an effective adjuvant is a substance that will stimulate the immune system nonspecifically.

- 5 These substances stimulate the production of a large set of soluble peptide factors known as lymphokines. In turn, lymphokines stimulate the activity of antigen-processing cells directly and cause a local inflammatory reaction at the site of injection. A component of lipopolysaccharide known as lipid A is commonly used. Lipid A is
10 available in a number of synthetic and natural forms that are much less toxic than lipopolysaccharides, but still retain most of the desirable adjuvant properties of the lipopolysaccharide molecules. Lipid A compounds are often delivered using liposomes. The two bacteria that are commonly used in adjuvants as non-specific stimulants are
15 Bordatella pertussis and Mycobacterium tuberculosis. When used as whole bacteria, they must be heat-killed prior to use. The immunomodulatory mediators of B. pertussis include a lipopolysaccharide component and the pertussis toxin. The pertussis toxin has been purified and is available commercially. M. tuberculosis
20 is commonly found in complete Freund's adjuvant. The most active component of M. tuberculosis has been localized to muramyl dipeptide which is available in a number of forms.

Immunizations (Inoculation and Booster Shots)

- 25 The subject to be immunized can be any mammal which possesses a competent immune system. Examples of subject mammals include humans and domestic animals (e.g. dogs, cats, cows, horses, etc.), as well as animals intended for experimental or other purposes (e.g., mice, rats, rabbits, etc.).
30 Two different criteria are important to consider in determining the proper dose for the initial immunization. First, the optimum dose to achieve the strongest response and second, the

- 13 -

minimum dose likely to induce the production of useful polyclonal antibodies. Much of the injected material will be catabolized and cleared before reaching the appropriate target immune cell. The efficiency of this process will vary with host factors, the route of injection, the use of adjuvants, and the intrinsic nature of the surface protein injected. Thus, the effective dose delivered to the immune system may bear little relationship to the introduced dose and consequently dose requirements must be determined empirically. These determinations can be readily made by one skilled in the art. Secondary injections and later boost can be given with amounts similar to or less than the primary injection.

The route of injection is guided by three practical decisions: 1) what volume must be delivered; 2) what buffers and other components will be injected with the immunogen; and 3) how quickly should the immunogen be released into the lymphatics or circulation. For example, with rabbits, large volume injections normally are given at multiple subcutaneous sites. For mice, large volumes are only possible with intraperitoneal injections. If adjuvants or particulate matter are included in the injection, the immunogen should not be delivered intravenously. If a slow release or the inoculant is desired, the injections should be done either intramuscularly or intradermally. For immediate release, use intravenous injections.

Primary antibody responses often are very weak, particularly for readily catabolized, soluble antigens. Hence, secondary or booster injections are required after the initial immunization. A delay is needed before reintroducing the protein into a primed subject. A minimum of 2 or 3 weeks is recommended but greater intervals are possible. The antibody responses to secondary and subsequent injections is much stronger. Higher titers of antibody are reached, but more importantly, the nature and quantity of the antibodies present in serum changes. These changes yield high-affinity antibodies. The intervals between secondary, tertiary and subsequent injections may also be

- 14 -

varied, but usually need to be extended to allow the circulating level of antibody to drop enough to prevent rapid clearance of newly injected antigen.

Subsequent booster injections will be required to increase
5 reduced circulating antibody for continued contraception. The actual intervals for these injections will differ from species to species. However, the intervals can be determined by one skilled in the art by monitoring serum levels of sperm surface protein antibodies.

In another embodiment, subjects can be administered with
10 alloantisera, or monoclonal antibodies, directed to a sperm surface protein to achieve contraception. The alloantiserum is raised in another individual of the same species, isolated from the serum of the individual and prepared in a suitable carrier for injection into the recipient subject. Those skilled in the art are familiar with methods for preparing and
15 formulating monoclonal antibodies for administration.

There is convincing evidence that naturally occurring antibodies to sperm cause infertility in women [Bronson, R.A., et al., *Fertility and Sterility*, 42: 171-183 (1984)]. This infertility is better correlated with the antibody titers in cervical mucus than with the serum
20 [Clark, G.N., *Amer. J Reprod. Immunol.*, 5:179-181 (1984)]. Presence of anti-sperm antibodies in the cervical mucus of infertile women results in poor sperm penetration through the cervical mucus and agglutination of the sperm, thereby reducing the number of sperm available for fertilization. Thus, success of a contraceptive vaccine depends in particular on the
25 generation of mucosal immune responses involving sustained titers of anti-sperm antibodies in the female reproductive tract.

Generally, local application of the antigen is an effective way to stimulate an antibody response by that mucosa [Mestecky, J., *J Clin. Immunol.*, 7: 265-276 (1987)]. However, local mucosal immunization is
30 ineffective in female reproductive tract due to the barrier function of the luminal epithelium and to rapid loss of antigen from the lumen of reproductive tract. Stability and adhesiveness of the antigen on the mucosal

- 15 -

- surface is important for the induction of the mucosal immune responses [de Aizpurua, H.J. and Russell-Jones, G.J., *J Exp. Med.*, 167: 440 (1988)]. Adhesive antigens are critical to successful mucosal immunization, not only because they are effective mucosal immunogens themselves, but also
- 5 because they are carrier proteins for other antigens. Cholera toxin is a potent immunogen when given mucosally, but acts as an adjuvant when given in combination with other antigens [McKenzie, S.J. and Halsey, J.A., *J. Immunol.*, 133: 1818 (1984)]. Effective immunization is also dependent on the stability of the antigen on a mucosal surface. Many antigens for use
- 10 in mucosal vaccines are poorly immunogenic because they are unable to survive in the acidic and proteolytic conditions of the mucosal surface [O'Hagen, D.T., *Curr. Opin. Infect. Dis.*, 3:393 (1990)]. The DL-lactide-co-glycolide (DL-PLG) microsphere, microparticle carrier system is one of the most suitable systems for mucosal immunization. DL-PLG
- 15 microspheres protect the antigen at mucosal surface and are taken up by the mucosal lymphoid tissues where they induce mucosal immunity [Eldridge, J.H. et al, *Curr. Top. Microbiol. Immunol.*, 146: 59 (1989)]. Liposomes and inactivated micro-organisms also are used as microparticle carriers. Some parenteral adjuvants such as Avridine, a lipoidal amine and muramyl
- 20 dipeptide (MDP), the active component of mycobacteria in Freund's complete adjuvant, also have been shown to be active as oral mucosal adjuvants and enhance mucosal immunization [Anderson, A.O. and Reynolds, J.A., *J. Reticuloendothel. Soc.*, 26(suppl): 667 (1979); Taubman, M.A., et al., *Ann. NY Acad. Sci.*, 409: 637 (1983)]. Development of
- 25 mucosal immune responses in female reproductive tract are optimized by using various adjuvants, micro particle carriers, by immunizing at local or remote mucosal surfaces or by combination of parenteral and mucosal immunization.
- 30 Utility of PH30 beta in Identification of Small Molecules that will Disrupt Sperm-egg Interaction and Fertilization

- 16 -

The comparison of the protein sequences of both mouse and human PH30 beta chain genes shows significant homology to a class of proteins called disintegrins found in the snake venoms. These proteins are known to bind a family of cell surface molecules called integrins and prevent their normal function in cell adhesion. On the basis of these homologies it is reasonable to conclude that the PH30 receptor on the oocyte is an integrin. Comparisons of the disintegrin domain sequences of guinea pig, mouse and human PH30 beta chain genes show significant differences in their putative ligand binding domain. In particular, the sequences in this region are different from other disintegrins and among the three species. The recombinant mouse and human PH30 beta proteins are used to make affinity resins to purify, identify and characterize mouse and human PH30 receptors. The recombinant PH30 beta also are used to determine its relative affinity to other integrins expressed in other tissues and are used as a ligand for cloning of the PH30 receptor.

Since the integrin recognition sequences in PH30 beta are species specific, the sequence information is necessary to identify small molecules that disrupt fertilization in a species specific manner. The recombinant mouse and human PH30 beta are used to set up screens to identify small molecules that act either as antagonist to PH30 receptor and disrupt PH30 binding or act as an agonist and stimulate PH30 receptor inducing transmembrane signaling, egg cortical granule release and zona reaction thus making the egg impenetrable for fertilization.

The present invention is further illustrated in the following exemplification.

EXAMPLE 1

Isolation of DNA Encoding Mouse and Human PH30 beta A. cDNA Library Plating

- 17 -

One million independent recombinant bacteriophage from both a human testis cDNA library in λ gt 11 (Clontech, Palo Alto, CA.) and mouse testis cDNA library (Stratagene La Jolla, CA.) in UNI-ZAP XR were plated. Plaque lifts were done in duplicate by placing a
5 nitrocellulose filter on the plate for two minutes, and treating the filter with denaturing solution (0.5M NaOH, 1.5M NaCl), neutralization buffer (0.5M Tris pH 7.5, 1.5M NaCl) and 2X SSC (3M NaCl, 0.35M sodium citrate pH 7.0) for two minutes each. The filters were dried for
10 thirty minutes at room temperature and then baked for two hours at 80°C in a vacuum oven.

B. Generation of Probe:

A guinea pig PH30 beta cDNA was isolated by RT-PCR (reverse transcriptase-polymerase chain reaction) as a 1020 bp (base
15 pairs), HindIII/Bam HI fragment, containing 94% of the coding sequence. This fragment was subcloned into pBluescript SK⁺ vector (Stratagene, La Jolla, CA) and verified by sequence analysis. A probe was made by nick translating the purified 1020 bp guinea pig PH30 beta fragment. The filters were probed at 42°C for fifteen hours in
20 hybridization solution (7mM Tris pH 7.5, 40% formamide, 4X SSC, 0.8X Denhard's, 20 μ g/ml of salmon sperm DNA and 10% Dextran sulfate) containing 10^6 cpm (counts per minute)/ml of the labeled probe. The filters were washed twice at room temperature for fifteen
25 minutes each with 2X SSC/0.2% SDS (sodium dodecyl sulfate), then twice at room temperature with 0.2X SSC/0.1%SDS, then once at 42°C for 30 minutes with 0.1X SSC/0.1%SDS. The filters were exposed to XAR film (Eastman Kodak Co, Rochester, NY) for 15 hours. The positive plaques were picked into 1 ml of SM (0.1M NaCl, 10mM Magnesium Sulphate, 2% gelatin, 50mM Tris pH 7.5) and screened
30 again as described above. After four rounds of screening, the purified plaques were obtained.

- 18 -

Purified plaques of mouse testicular library were subcloned into pBluescript SK⁺ vector using the EX ASSIT helper phage and SOLR cells (Stratagene, La Jolla, CA). DNA from the purified plaques of human testicular library was isolated using light PLG 2 tubes and following manufacturer's (Clontech, Palo Alto, CA) directions. The DNA was then digested with the restriction enzyme EcoRI and ligated into pBluescript SK⁺ and was used to transform competent *E. coli* strain HB101 cells.

10 C. DNA Sequencing and Analysis:

Cloned inserts were sequenced on both strands using the Sequenase kit (United States Biochemical, Cleveland, Ohio). Sequences were analyzed by searching GeneBank and EMBL DNA sequence database using the FASTA program (University of Wisconsin, Genetics Computer Group) and sequence comparisons were done using the GAP program.

D. Characterization of cDNA Clones:

The screening of the mouse testicular library with a 1020 bp guinea pig PH30 beta probe resulted in the isolation of a 1.7 kb (kilo base pair) cDNA clone. This cDNA clone contains a 1371 nucleotide open reading frame and a 329 nucleotide 3' untranslated region. When mature parts of the guinea pig and mouse PH30 beta were compared, the mouse PH30 beta clone showed a maximum of 63% identity to guinea pig PH30 beta at the nucleotide level. The amino terminal 103 residues of the deduced 457 amino acid sequence represents the precursor regions of the mouse PH30 beta that are cleaved off at sperm maturation. At the amino acid level the mature mouse, and guinea pig PH30 betas were 54% identical with all the cysteines lining up.

30 The human testicular cDNA library screening identified a 2.331 kb cDNA which contains an open reading frame of 1959 nucleotides and 372 nucleotide 3' untranslated region. The human PH30

- 19 -

beta clone was 63 and 67% identical in its open reading frame to mouse and guinea pig PH30 beta genes, respectively. Comparison of the derived 653 amino acid sequence with the mouse and guinea pig PH30 beta indicates that the amino terminal 299 represents the precursor and
5 carboxy terminal 354 amino acids represent the mature part of human PH30 beta respectively. The amino acid sequence of the mature human PH30 beta was 54% homologous to mature guinea pig and mouse PH30 beta proteins.

Protein sequence comparison of mouse and human PH30
10 beta to guinea pig PH30 beta and snake venom disintegrins indicated significant homology. This analysis revealed similar structural organization and indicated the presence of metalloprotease and disintegrin domains in these proteins.

Metalloprotease domains of mouse and human PH30 beta
15 shared significant similarity with the metalloprotease domains of guinea pig PH30 beta but less similarity to the metalloprotease domain of guinea pig PH30 alpha or other disintegrins. The active site signature sequence of zinc-dependent metalloproteases is present in PH30 alpha and the snake venom disintegrins, Jararhagin and Trigramin.
20 [Wolfsberg, T.G., et al., Proc. Natl. Acad. Sci. USA 90: 10783-10797 (1993)]. Similar to guinea pig PH30 beta, the mouse and human metalloprotease domain lacks the active site signature sequence and both were 80% identical to guinea pig PH30 beta and only 30% identical to guinea pig PH30 alpha metalloprotease active site sequence. Human and
25 guinea pig PH30 beta metalloprotease domains were 60% identical.

Similar to guinea pig PH30 beta, the mouse and human PH30 beta also contain a disintegrin domain. The disintegrin domain in mouse PH30 beta contains 91 amino acids (residues 111-202) and in human, 93 amino acids (residues 299-392). Most disintegrins of snake
30 venom contain a consensus integrin binding sequence RGD. Another family of snake venom disintegrins that are linked to a carboxyl terminus cysteine rich domain, lack the RGD tripeptide but contain a

- 20 -

unique tripeptide and adjacent cysteine. Guinea pig, mouse and human PH30 beta proteins also do not contain RGD tripeptide and share more similarity with this later family of disintegrins. These snake venom disintegrins and disintegrin domains of guinea pig, mouse and human PH30 beta contain a negatively charged residue at the carboxyl end of the tripeptide sequence. The integrin binding sequence of guinea pig PH30 beta is TDE. One skilled in the art would have expected that the integrin binding site of PH30 beta of other mammalian species would also be TDE. However, after isolation of human and mouse PH30 beta, it was found that this was not the case. It was unexpectedly discovered that the critical sequence at the integrin binding site was not conserved. Comparisons of guinea pig, mouse and human PH30 beta disintegrin domains showed significant variation in their putative integrin binding sequences although the carboxy terminus end of these domains were identical. The putative integrin binding residues in PH30 beta were QDE in mouse and FEE in human. These differences in the integrin binding sequences between species were an unexpected and surprising finding.

Both mouse and human PH30 beta contain an epidermal growth factor like repeat and a transmembrane domain that are 60% identical to similar regions of guinea pig PH30 beta.

EXAMPLE 2

25 Cloning of the 5' end of Mouse and Human PH30 Beta

The 5' ends of mouse and human PH30 beta were cloned using the Gibco BRL "5' RACE System for Rapid Amplification of cDNA Ends" and following manufacturer's protocols. 2 oligonucleotides were synthesized for each template. Oligo 1 was an antisense primer and Oligo 2 was also an antisense primer, internal to oligo 1, and contained in the CAU sequences on the 5' end to facilitate cloning. Oligo 1 was annealed to mouse or human testis mRNA and a

- 21 -

cDNA copy was made using SuperScript II Reverse Transcriptase. The mRNA template was degraded with Rnase H. The single strands cDNA copy was purified with GlassMAX Spin columns and was then tailed on the 3' end with dCTP and terminal transferase. The tailed cDNA copy
5 was then amplified using a supplied anchor primer that contains the 5'CAU cloning site and oligo 2. The amplification system was Taq polymerase. The amplified product was then gel purified, treated with Uracil DNA Glycosylase, subcloned into the vector pAMP1 and then transformed into competent E. coli DH5 cells. Colonies were identified
10 which had subcloned fragment and these colonies were sequenced as described previously.

The complete mouse cDNA sequence and the deduced amino acid sequence of the mouse PH30 beta protein is shown in SEQ ID NO: 5 and SEQ ID NO: 6. The complete human cDNA
15 sequence and the deduced amino acid sequence of the human PH30 beta protein is shown in SEQ ID NO: 7 and SEQ ID NO: 8.

At the nucleotide level, the complete human PH30 beta shares 68% identity with mouse and 68.6% identity with guinea pig PH30 beta, respectively. Mouse and guinea pig DNA sequences are
20 65.5% identical. The amino acid sequence of the human PH30 beta is 58.9% identical to mouse and 56.5% identical to guinea pig PH30 beta. At the amino acid level, the mouse and guinea pig PH30 beta are 55.2% identical.

EXAMPLE 3

25

Contraceptive Vaccination by the Administration of PH30 beta Protein

Female or male mice (about 7 weeks old at the time of first injection) receive two injections of PH30 beta in the amounts stated below. Recombinant or native PH30 beta, purified from cell line or
30 sperm by mAb-affinity chromatography or biochemical methods, shows at least 90% purity (i.e., no more than 10% detectable contaminants) using silver-staining of purified protein on SDS gels. Purity of each

- 22 -

PH30 preparation used for immunization of females or males is verified by SDS polyacrylamide gel electrophoresis and silver staining. The affinity-purified PH30 beta, in 0.375 ml phosphate-buffered saline (PBS) containing 3 mM octylglucoside (OG) is emulsified with 0.375 ml complete Freund's adjuvant (CFA). Each animal receives 0.1 ml of the emulsion subcutaneously in the back and 0.05 ml intramuscularly in a rear leg. About 3 weeks later, the same amount of PH30 beta in PBS and 3 mM OG is emulsified with incomplete Freund's adjuvant (IFA), and is injected in the same sites in each animal. Control females and males receive the same injections on the same schedule and containing PBS and 3 mM OG and CFA or IFA, but lacking PH30 beta. To allow the injected females to mate, about 6 weeks after the initial injection they are housed with males for 10 days. Each cage contains one male (13 weeks old), one PH30 beta immunized female, and from 2-4 control injected females. 24 hours after the grouping, females are checked visually daily for the vaginal plugs. Two weeks after the initiation of the mating the, females are removed into individual cages. After three weeks the pregnant females having litters and progeny are counted. To allow the injected males to mate, about six weeks after the initial injection, each injected male is housed with two females (10-13 weeks) for 10 days. The females and males are then separated and after an additional 3 weeks pups are counted.

EXAMPLE 4

25

Use of PH30 Disintegrin Peptides as Inhibitor of Sperm Fusion to Egg Plasma Membrane

Peptides from the PH30 β disintegrin domain are tested for inhibition of sperm binding to the egg plasma membrane.

30

The fusion inhibition assay is carried out as follows. Young female mice (8-10 weeks of age) are injected with 5 units of pregnant mare's serum (PMS) in 0.9 NaCl intraperitoneally. 48 hours later, the

- 23 -

mice are injected IP with 5 units of hCG (human chorionic gonadotrophin) in 0.9% NaCl to trigger super ovulation. 14-16 hours after hCG injection, the ovulated oocytes are collected and treated with hyaluronidase to remove cumulus cells. The zona pellucida is removed with a mixture of proteases. The zona pellucida free eggs are incubated in culture media with peptide at a specified concentration for 30 minutes [Hogan, B. et al., *Manipulating The Mouse Embryo*, 91-101, (1986)]. Sperm collected from the epididymis of male mice is capacitated by incubation and acrosome reacted as described by Fleming and Yanagimachi [*Gamete Res.* 4, 253-273 (1981)] and added to the eggs and incubated for 15 minutes. The eggs are then transferred to a sperm free culture medium and incubated for an additional 1 hour and 45 minutes. The eggs are then fixed and stained as described by Primakoff et al., [*J. Cell. Biol.* 104, 141 (1987)]. The total number of swollen sperm heads are then counted. Swollen sperm heads are an indication that the sperm and egg have fused.

On the basis of these observations, several indices are calculated. The fertilization index (F.I.) is determined by dividing the total number of swollen heads by the total number of eggs. The fertilization rate (F.R.) is the percentage of eggs fertilized. The percent inhibition is determined by dividing the fertilization index of the experimental peptide by the fertilization index of the control peptide.

The PH30 β disintegrin domain represents an epitope which is critical in sperm-egg fusion. Antibodies which bind specifically to this epitope block sperm/egg fusion.

EXAMPLE 5

Use of PH30 beta to Identify Small Molecules that will disrupt sperm-egg Interaction and Fertilization

A. Identification of PH30 beta receptor antagonists:

- 24 -

Identification of compounds that specifically interfere with the binding of PH30 to their receptor on the egg, has been limited due to unavailability of the sufficient quantities of PH30 protein and normal human eggs. The availability of the rPH30 beta facilitates the identification and cloning of PH30 beta receptor integrin cDNAs. These PH30 beta receptor cDNAs are used to generate recombinant PH30 beta receptors. The alternative source of PH30 beta receptors facilitates identification of substances that affect the binding of PH30 beta to its receptors.

Using conventional methods, the Chinese Hamster Ovary cells are transfected with cDNAs encoding the PH30 beta receptor to produce a stable transformed cell which expresses human PH30 beta receptor integrin in large quantities. Such a transformed cell provides a consistent source of recombinant PH30 beta receptors and is useful in the characterization of the binding of PH30 beta to its receptor and for establishing assays to screen for compounds that inhibit PH30 binding to its receptor.

Selectivity of the compounds to PH30 beta receptor is examined by using cell lines that express other integrin receptors that contain the same beta subunit and closely related alpha chain. Compounds that specifically inhibit PH30 beta/receptor interaction are tested further in biological assays, like inhibition of sperm-egg fusion assay and egg cortical granule release assay to determine their efficacy in inhibiting fertilization.

B. Protocol for PH30 beta antagonist screen:

Cells expressing PH30 beta receptor are treated with extraction buffer (50 mM Tris pH 7.6, 100 mM n-Octyl β -D-Glucopyranoside, 150 mM NaCl, 1 mM MgCl₂ and 1 mM CaCl₂) and soluble material is separated by centrifugation and stored frozen at -80 °C. In an assay tube the 15 μ l water, 80 μ l of assay buffer (125 mM Tris pH 7.6, 187.5 mM NaCl, 1.25 mM CaCl₂, 1.25 mM MgCl₂ and 1.25% BSA) and 5 μ l of sample compound or control (40 μ M of cold PH30 beta) are added and mixed with 50 μ l of ¹²⁵I-PH30 beta (final concentration 40 pM) and 50 μ l of cell extract (final protein concentration 250 μ g/ml). The tubes

- 25 -

are incubated at room temperature for 1 hour. Following incubation the samples are harvested using Tomtec Mach II- 6x 16 cell harvester and printed filtermat cat. # 1205-404. Filters are dried and counted in LKB/Wallac Beta Plate counter.

5

Calculations and Interpretations:

$$\% \text{ Inhibition} = \frac{\text{CPMavg total binding} - \text{CPMavg sample}}{\text{CPMavg total binding} - \text{CPMavg positive control}} \times 100$$

10

When % inhibition > 60 and the inhibition is dose related, the sample will be considered active.

C. Sperm-Oocyte fusion assay:

15

Young female mice (approximately 8-10 weeks of age) are injected with 5 units of pregnant mare's serum (PMS) in 0.9 NaCl intraperitoneally. 48 hours later, the mice are injected IP with 5 units of hCG (human chorionic gonadotrophin) in 0.9% NaCl to trigger super ovulation. 14-16 hours after hCG injection, the ovulated oocytes are collected and treated with hyaluronidase to remove cumulus cells. Zona pellucida is removed by treating eggs briefly with 0.1 mg/ml of chymotrypsin. Oocytes are washed with Hepes buffered culture medium and are loaded with a fluorescent stain 4',6-diamidino-2-phenylindole dihydrochloride (DAPI) by incubating at 37°C for 30 minutes. Oocytes are then washed with medium and incubated with rPH30 beta or inhibitor compound for 30 minutes followed by another 30 minute incubation with 1×10^4 sperms that have been previously capacitated by incubating with calcium ionophore. After incubation, the oocytes are washed, mounted and examined by light microscopy and scored for the presence of fluorescent swollen sperm heads with associated tails in cytoplasm.

20

25

30

- 26 -

Fertilization rate = $\frac{\text{number of eggs fused}}{\text{number of eggs tested}} \times 100$ (results expressed as % fertilization)

5 In the absence of any inhibitor > 90% oocytes are fertilized. When the sperm-oocyte fusion is inhibited >60% and the inhibition is dose related the compound will be considered active.

10 While the invention has been described and illustrated with reference to certain preferred embodiments thereof, those skilled in the art will appreciate that various changes, modifications and substitutions can be made therein without departing from the spirit and scope of the invention. It is intended, therefore, that the invention be limited only by the scope of the claims which follow and that such claims be interpreted as broadly as is reasonable.

- 27 -

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: ALVES, KENNETH
GUPTA, SUNIL K.
HOLLIS, GREGORY F.
- (ii) TITLE OF INVENTION: CONTRACEPTIVE VACCINE
- (iii) NUMBER OF SEQUENCES: 8
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: MARY A. APPOLLINA
 - (B) STREET: P.O. BOX 2000, 126 E. LINCOLN AVENUE
 - (C) CITY: RAHWAY
 - (D) STATE: NJ
 - (E) COUNTRY: USA
 - (F) ZIP: 07065
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: US
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: APPOLLINA, MARY A
 - (B) REGISTRATION NUMBER: 34,087
 - (C) REFERENCE/DOCKET NUMBER: 19244Y
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (908)594-3462
 - (B) TELEFAX: (908)594-4720

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 2373 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

- 28 -

GGCCAAGATT TTCAGAATTT CTGCCACTAC CAAGGGTATA TTGAAGGTTA TCCAAAATCT	60
GTGGTGATGG TTAGCACATG TACTGGACTC AGGGGCGTAC TACAGTTTGA AAATGTTAGT	120
TATGGAATAG AACCCCTGGA GTCTTCAGTT GGCTTTGAAC ATGTAATTTA CCAAGTAAAA	180
CATAAGAAAG CAGATGTTTC CTTATATAAT GAGAAGGATA TTGAATCAAG AGATCTGTCC	240
TTTAAATTAC AAAGCGCAGA GCCACAGCAA GATTTTGCAA AGTATATAGA AATGCATGTT	300
ATAGTTGAAA AACAATTGTA TAATCATATG GGGTCTGATA CAACTGTTGT CGCTCAAAAA	360
GTTTTCCAGT TGATTGGATT GACGAATGCT ATTTTTGTTT CATTTAATAT TACAATTATT	420
CTGTCTTCAT TGGAGCTTTG GATAGATGAA AATAAAATTG CAACCACTGG AGAAGCTAAT	480
GAGTTATTAC ACACATTTTT AAGATGGAAA ACATCTTATC TTGTTTTACG TCCTCATGAT	540
GTGGCATTTT TACTTGTTTA CAGAGAAAAG TCAAATTATG TTGGTGCAAC CTTTCAAGGG	600
AAGATGTGTG ATGCAAACTA TGCAGGAGGT GTTGTCTGTC ACCCCAGAAC CATAAGTCTG	660
GAATCACTTG CAGTTATTTT AGCTCAATTA TTGAGCCTTA GTATGGGGAT CACTTATGAT	720
GACATTAACA AATGCCAGTG CTCAGGAGCT GTCTGCATTA TGAATCCAGA AGCAATTCAT	780
TTCAGTGGTG TGAAGATCTT TAGTAACTGC AGCTTCGAAG ACTTTGCACA TTTTATTTCA	840
AAGCAGAAAT CCCAGTGTCT TCACAATCAG CCTCGCTTAG ATCCTTTTTT CAAACAGCAA	900
GCAGTGTGTG GTAATGCAAA GCTGGAAGCA GGAGAGGAGT GTGACTGTGG GACTGAACAG	960
GATTGTGCCC TTATTGGAGA AACATGCTGT GATATTGCCA CATGTAGATT TAAAGCCGGT	1020
TCAAACCTTG CTGAAGGACC ATGCTGCGAA AACTGTCTAT TTATGTCAA AGAAAGAATG	1080
TGTAGGCCTT CCTTTGAAGA ATGCGACCTC CCTGAATATT GCAATGGATC ATCTGCATCA	1140
TGCCCAGAAA ACCACTATGT TCAGACTGGG CATCCGTGTG GACTGAATCA ATGGATCTGT	1200
ATAGATGGAG TTTGTATGAG TGGGGATAAA CAATGTACAG ACACATTTGG CAAAGAAGTA	1260
GAGTTTGGCC CTTCAGAATG TTATTCTCAC CTTAATTCAA AGACTGATGT ATCTGGAAC	1320
TGTGGTATAA GTGATTCAGG ATACACACAG TGTGAAGCTG ACAATCTGCA GTGCGGAAAA	1380
TTAATATGTA AATATGTAGG TAAATTTTTA TTACAAATTC CAAGAGCCAC TATTATTTAT	1440
GCCAACATAA GTGGACATCT CTGCATTGCT GTGGAATTTG CCAGTGATCA TGCAGACAGC	1500
CAAAAGATGT GGATAAAGA TGGAACCTCT TGTGGTTCAA ATAAGGTTTG CAGGAATCAA	1560
AGATGTGTGA GTTCTTCATA CTTGGGTTAT GATTGTACTA CTGACAAATG CAATGATAGA	1620
GGTGTATGCA ATAACAAAAA GCACTGTCAC TGTAGTGCTT CATATTTACC TCCAGATTGC	1680

- 29 -

TCAGTTCAAT CAGATCTATG GCCTGGTGGG AGTATTGACA GTGGCAATTT TCCACCTGTA 1740
 GCTATACCAG CCAGACTCCC TGAAAGGCGC TACATTGAGA ACATTTACCA TTCCAAACCA 1800
 ATGAGATGGC CATTTTCTT ATTCATTCTT TTCTTTATTA TTTTCTGTGT ACTGATTGCT 1860
 ATAATGGTGA AAGTTAATTT CCAAAGGAAA AAATGGAGAA CTGAGGACTA TTCAAGCGAT 1920
 GAGCAACCTG AAAGTGAGAG TGAACCTAAA GGGTAGTCTG GACAACAGAG ATGCCATGAT 1980
 ATCACTTCTT CTAGAGTAAT TATCTGTGAT GGATGGACAC AAAAAAATGG AAAGAAAAGA 2040
 ATGTACATTA CCTGGTTTCC TGGGATTCAA ACCTGCATAT TGTGATTTTA ATTTGACCAG 2100
 AAAATATGAT ATATATGTAT AATTTCACAG ATAATTTACT TATTTAAAAA TGCATGATAA 2160
 TGAGTTTTAC ATTACAAATT TCTGTTTTTT TAAAGTTATC TTACGCTATT TCTGTTGGTT 2220
 AGTAGACACT AATTCGTGCA GTAGGGGCAT GGTATAAGGA AATATCATAA TGTAATGAGG 2280
 TGGTACTATG ATTAAAAGCC ACTGTTACAT TTCAAAAAAA AAAAAAAAAA ACCATCTAAA 2340
 AAAGGTAGGT AGGTAAAAGA ATTATATTAT CAA 2373

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 651 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Gly Gln Asp Phe Gln Asn Phe Cys His Tyr Gln Gly Tyr Ile Glu Gly
 1 5 10 15
 Tyr Pro Lys Ser Val Val Met Val Ser Thr Cys Thr Gly Leu Arg Gly
 20 25 30
 Val Leu Gln Phe Glu Asn Val Ser Tyr Gly Ile Glu Pro Leu Glu Ser
 35 40 45
 Ser Val Gly Phe Glu His Val Ile Tyr Gln Val Lys His Lys Lys Ala
 50 55 60
 Asp Val Ser Leu Tyr Asn Glu Lys Asp Ile Glu Ser Arg Asp Leu Ser
 65 70 75 80
 Phe Lys Leu Gln Ser Ala Glu Pro Gln Gln Asp Phe Ala Lys Tyr Ile
 85 90 95

- 30 -

Glu Met His Val Ile Val Glu Lys Gln Leu Tyr Asn His Met Gly Ser	100	105	110
Asp Thr Thr Val Val Ala Gln Lys Val Phe Gln Leu Ile Gly Leu Thr	115	120	125
Asn Ala Ile Phe Val Ser Phe Asn Ile Thr Ile Ile Leu Ser Ser Leu	130	135	140
Glu Leu Trp Ile Asp Glu Asn Lys Ile Ala Thr Thr Gly Glu Ala Asn	145	150	155
Glu Leu Leu His Thr Phe Leu Arg Trp Lys Thr Ser Tyr Leu Val Leu	165	170	175
Arg Pro His Asp Val Ala Phe Leu Leu Val Tyr Arg Glu Lys Ser Asn	180	185	190
Tyr Val Gly Ala Thr Phe Gln Gly Lys Met Cys Asp Ala Asn Tyr Ala	195	200	205
Gly Gly Val Val Leu His Pro Arg Thr Ile Ser Leu Glu Ser Leu Ala	210	215	220
Val Ile Leu Ala Gln Leu Leu Ser Leu Ser Met Gly Ile Thr Tyr Asp	225	230	235
Asp Ile Asn Lys Cys Gln Cys Ser Gly Ala Val Cys Ile Met Asn Pro	245	250	255
Glu Ala Ile His Phe Ser Gly Val Lys Ile Phe Ser Asn Cys Ser Phe	260	265	270
Glu Asp Phe Ala His Phe Ile Ser Lys Gln Lys Ser Gln Cys Leu His	275	280	285
Asn Gln Pro Arg Leu Asp Pro Phe Phe Lys Gln Gln Ala Val Cys Gly	290	295	300
Asn Ala Lys Leu Glu Ala Gly Glu Glu Cys Asp Cys Gly Thr Glu Gln	305	310	315
Asp Cys Ala Leu Ile Gly Glu Thr Cys Cys Asp Ile Ala Thr Cys Arg	325	330	335
Phe Lys Ala Gly Ser Asn Cys Ala Glu Gly Pro Cys Cys Glu Asn Cys	340	345	350
Leu Phe Met Ser Lys Glu Arg Met Cys Arg Pro Ser Phe Glu Glu Cys	355	360	365
Asp Leu Pro Glu Tyr Cys Asn Gly Ser Ser Ala Ser Cys Pro Glu Asn	370	375	380
His Tyr Val Gln Thr Gly His Pro Cys Gly Leu Asn Gln Trp Ile Cys	385	390	395
			400

- 31 -

Ile Asp Gly Val Cys Met Ser Gly Asp Lys Gln Cys Thr Asp Thr Phe
 405 410 415
 Gly Lys Glu Val Glu Phe Gly Pro Ser Glu Cys Tyr Ser His Leu Asn
 420 425 430
 Ser Lys Thr Asp Val Ser Gly Asn Cys Gly Ile Ser Asp Ser Gly Tyr
 435 440 445
 Thr Gln Cys Glu Ala Asp Asn Leu Gln Cys Gly Lys Leu Ile Cys Lys
 450 455 460
 Tyr Val Gly Lys Phe Leu Leu Gln Ile Pro Arg Ala Thr Ile Ile Tyr
 465 470 475 480
 Ala Asn Ile Ser Gly His Leu Cys Ile Ala Val Glu Phe Ala Ser Asp
 485 490 495
 His Ala Asp Ser Gln Lys Met Trp Ile Lys Asp Gly Thr Ser Cys Gly
 500 505 510
 Ser Asn Lys Val Cys Arg Asn Gln Arg Cys Val Ser Ser Ser Tyr Leu
 515 520 525
 Gly Tyr Asp Cys Thr Thr Asp Lys Cys Asn Asp Arg Gly Val Cys Asn
 530 535 540
 Asn Lys Lys His Cys His Cys Ser Ala Ser Tyr Leu Pro Pro Asp Cys
 545 550 555 560
 Ser Val Gln Ser Asp Leu Trp Pro Gly Gly Ser Ile Asp Ser Gly Asn
 565 570 575
 Phe Pro Pro Val Ala Ile Pro Ala Arg Leu Pro Glu Arg Arg Tyr Ile
 580 585 590
 Glu Asn Ile Tyr His Ser Lys Pro Met Arg Trp Pro Phe Phe Leu Phe
 595 600 605
 Ile Pro Phe Phe Ile Ile Phe Cys Val Leu Ile Ala Ile Met Val Lys
 610 615 620
 Val Asn Phe Gln Arg Lys Lys Trp Arg Thr Glu Asp Tyr Ser Ser Asp
 625 630 635 640
 Glu Gln Pro Glu Ser Glu Ser Glu Pro Lys Gly
 645 650

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1768 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

- 32 -

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GGCACGAGCG ATTATGTTGG CGCTACCTAT CAAGGGAAGA TGTGTGACAA GAACTATGCA	60
GGAGGAGTTG CTTTGCACCC CAAAGCCGTA ACTCTGGAAT CACTTGCAAT TATTTTAGTT	120
CAGCTGCTGA GCCTCAGCAT GGGGCTAGCG TATGACGACG TGAACAAGTG CCAGTGTGGC	180
GTACCTGTCT GCGTGATGAA CCCGGAAGCG CCTCACTCCA GCGGTGTCCG GGCCTTCAGT	240
AACTGCAGCA TGGAGGACTT TTCCAAGTTT ATCACAAGTC AAAGCTCCCA CTGTCTGCAG	300
AACCAGCCAA CGCTACAGCC ATCTTACAAG ATGGCGGTCT GTGGGAATGG AGAGGTGGAA	360
GAAGATGAAA TTTGCGACTG TGGAAAGAAG GGCTGTGCAG AAATGCCCCC GCCATGCTGT	420
AACCCCGACA CCTGTAAGCT GTCAGATGGC TCCGAGTGCT CCAGCGGGAT ATGCTGCAAC	480
TCGTGCAAGC TGAAGCGGAA AGGGGAGGTT TGCAGGCTTG CCCAAGATGA GTGTGATGTC	540
ACAGAGTACT GCAACGGCAC ATCCGAAGTG TGTGAAGACT TCTTTGTTCA AAACGGTCAC	600
CCATGTGACA ATCGCAAGTG GATCTGTATT AACGGCACCT GTCAGAGTGG AGAACAGCAG	660
TGCCAGGATC TATTTGGCAT CGATGCAGGC TTTGGTTCAA GTGAATGTTT CTGGGAGCTG	720
AATTCCAAGA GCGACATATC TGGGAGCTGT GGAATCTCTG CTGGGGGATA CAAGGAATGC	780
CCACCTAATG ACCGGATGTG TGGGAAAATA ATATGTAAAT ACCAAAGTGA AAATATACTA	840
AAATTGAGGT CTGCCACTGT TATTTATGCC AATATAAGCG GGCATGTCTG CGTTTCCCTG	900
GAATATCCCC AAGGTCATAA TGAGAGCCAG AAGATGTGGG TGAGAGATGG AACCGTCTGC	960
GGGTCAATA AGGTTTGCCA GAATCAAAAA TGTGTAGCAG ACACTTTCTT GGGCTATGAT	1020
TGCAACCTGG AAAAATGCAA CCACCATGGT GTATGTAATA ACAAGAAGAA CTGCCACTGT	1080
GACCCACAT ACTTACCTCC AGATTGTAAG AGAATGAAAG ATTCATATCC TGGCGGGAGC	1140
ATTGATAGTG GCAACAAGGA AAGGGCTGAA CCCATCCCTG TACGGCCCTA CATTGCAAGT	1200
CGTTACCGCT CCAAGTCTCC ACGGTGGCCA TTTTCTTGA TCATCCCTTT CTACGTTGTG	1260
ATCCTTGTCC TGATTGGGAT GCTGGTAAAA GTCTATTCCC AAAGGATGAA ATGGAGAATG	1320
GATGACTTCT CAAGCGAAGA GCAATTTGAA AGTGAAAGTG AATCCAAAGA CTAGTCTGGA	1380
CAGATTCCAC AATGTCACAA GTAATTCTCT TCAGTGACAA GAAAAAAAAG TGGAAAAGAA	1440
AAGCCTATGC ATTATCTTGC CTGAAAGTCA AGCCTGCATA TCGTGGTCTC CATCAGGCCA	1500

- 33 -

GAAATCATAT CTCTCCATTA CACATGTATG ATACATATGT GTGTATATTA TTCCATAAAT 1560
 GATTTACTTG TAAGAAATGA ATGATTATGA ATTTTCATATT ATACTTTGAT ATTTTACCCT 1620
 ATTTCTGGTA GTCGGTAGTC ATCAATTGTA TTTTCTAGTA GGTACATTAT AGAAAAGGCT 1680
 ATAAGAAAAT AAATGTGGTA CCATAATAAT CAATATCATA CAACCACCAT CTAAAAAAGG 1740
 TAGGTAGGTA AAAGAATTAT ATTATCAA 1768

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 457 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Gly	Thr	Ser	Asp	Tyr	Val	Gly	Ala	Thr	Tyr	Gln	Gly	Lys	Met	Cys	Asp	1	5	10	15
Lys	Asn	Tyr	Ala	Gly	Gly	Val	Ala	Leu	His	Pro	Lys	Ala	Val	Thr	Leu	20	25	30	
Glu	Ser	Leu	Ala	Ile	Ile	Leu	Val	Gln	Leu	Leu	Ser	Leu	Ser	Met	Gly	35	40	45	
Leu	Ala	Tyr	Asp	Asp	Val	Asn	Lys	Cys	Gln	Cys	Gly	Val	Pro	Val	Cys	50	55	60	
Val	Met	Asn	Pro	Glu	Ala	Pro	His	Ser	Ser	Gly	Val	Arg	Ala	Phe	Ser	65	70	75	80
Asn	Cys	Ser	Met	Glu	Asp	Phe	Ser	Lys	Phe	Ile	Thr	Ser	Gln	Ser	Ser	85	90	95	
His	Cys	Leu	Gln	Asn	Gln	Pro	Thr	Leu	Gln	Pro	Ser	Tyr	Lys	Met	Ala	100	105	110	
Val	Cys	Gly	Asn	Gly	Glu	Val	Glu	Glu	Asp	Glu	Ile	Cys	Asp	Cys	Gly	115	120	125	
Lys	Lys	Gly	Cys	Ala	Glu	Met	Pro	Pro	Pro	Cys	Cys	Asn	Pro	Asp	Thr	130	135	140	
Cys	Lys	Leu	Ser	Asp	Gly	Ser	Glu	Cys	Ser	Ser	Gly	Ile	Cys	Cys	Asn	145	150	155	160
Ser	Cys	Lys	Leu	Lys	Arg	Lys	Gly	Glu	Val	Cys	Arg	Leu	Ala	Gln	Asp				

- 34 -

	165		170		175
Glu Cys Asp Val Thr Glu Tyr Cys Asn Gly Thr Ser Glu Val Cys Glu	180		185		190
Asp Phe Phe Val Gln Asn Gly His Pro Cys Asp Asn Arg Lys Trp Ile	195		200		205
Cys Ile Asn Gly Thr Cys Gln Ser Gly Glu Gln Gln Cys Gln Asp Leu	210		215		220
Phe Gly Ile Asp Ala Gly Phe Gly Ser Ser Glu Cys Phe Trp Glu Leu	225		230		235
Asn Ser Lys Ser Asp Ile Ser Gly Ser Cys Gly Ile Ser Ala Gly Gly	245		250		255
Tyr Lys Glu Cys Pro Pro Asn Asp Arg Met Cys Gly Lys Ile Ile Cys	260		265		270
Lys Tyr Gln Ser Glu Asn Ile Leu Lys Leu Arg Ser Ala Thr Val Ile	275		280		285
Tyr Ala Asn Ile Ser Gly His Val Cys Val Ser Leu Glu Tyr Pro Gln	290		295		300
Gly His Asn Glu Ser Gln Lys Met Trp Val Arg Asp Gly Thr Val Cys	305		310		315
Gly Ser Asn Lys Val Cys Gln Asn Gln Lys Cys Val Ala Asp Thr Phe	325		330		335
Leu Gly Tyr Asp Cys Asn Leu Glu Lys Cys Asn His His Gly Val Cys	340		345		350
Asn Asn Lys Lys Asn Cys His Cys Asp Pro Thr Tyr Leu Pro Pro Asp	355		360		365
Cys Lys Arg Met Lys Asp Ser Tyr Pro Gly Gly Ser Ile Asp Ser Gly	370		375		380
Asn Lys Glu Arg Ala Glu Pro Ile Pro Val Arg Pro Tyr Ile Ala Ser	385		390		395
Arg Tyr Arg Ser Lys Ser Pro Arg Trp Pro Phe Phe Leu Ile Ile Pro	405		410		415
Phe Tyr Val Val Ile Leu Val Leu Ile Gly Met Leu Val Lys Val Tyr	420		425		430
Ser Gln Arg Met Lys Trp Arg Met Asp Asp Phe Ser Ser Glu Glu Gln	435		440		445
Phe Glu Ser Glu Ser Glu Ser Lys Asp	450		455		

- 35 -

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2553 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 17..2221

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

TGAGGAGGAC CAGCGC ATG CGG CTC ATC TTG CTT CTA CTG AGT GGG CTG	49
Met Arg Leu Ile Leu Leu Leu Leu Ser Gly Leu	
1 5 10	
AGT GAA CTT GGC GGC CTT AGC CAG TCC CAA ACA GAA GGC ACT CGT GAG	97
Ser Glu Leu Gly Gly Leu Ser Gln Ser Gln Thr Glu Gly Thr Arg Glu	
15 20 25	
AAA TTA CAC GTG CAA GTC ACA GTG CCA GAG AAA ATC CGG TCC GTC ACA	145
Lys Leu His Val Gln Val Thr Val Pro Glu Lys Ile Arg Ser Val Thr	
30 35 40	
AGC AAT GGC TAC GAA ACA CAG GTG ACC TAC AAT CTC AAA ATC GAA GGG	193
Ser Asn Gly Tyr Glu Thr Gln Val Thr Tyr Asn Leu Lys Ile Glu Gly	
45 50 55	
AAA ACA TAC ACC TTG GAC CTA ATG CAA AAA CCG TTC TTG CCT CCC AAC	241
Lys Thr Tyr Thr Leu Asp Leu Met Gln Lys Pro Phe Leu Pro Pro Asn	
60 65 70 75	
TTT AGA GTA TAC AGT TAT GAC AAC GCA GGA ATC ATG AGG TCT CTT GAG	289
Phe Arg Val Tyr Ser Tyr Asp Asn Ala Gly Ile Met Arg Ser Leu Glu	
80 85 90	
CAG AAG TTT CAG AAT ATC TGC TAC TTC CAA GGA TAC ATT GAA GGT TAT	337
Gln Lys Phe Gln Asn Ile Cys Tyr Phe Gln Gly Tyr Ile Glu Gly Tyr	
95 100 105	
CCA AAT TCT ATG GTG ATT GTT AGC ACA TGT ACT GGA CTG AGG GGT TTT	385
Pro Asn Ser Met Val Ile Val Ser Thr Cys Thr Gly Leu Arg Gly Phe	
110 115 120	
CTC CAA TTT GGA AAC GTT AGC TAT GGA ATT GAA CCT CTG GAA TCT TCC	433
Leu Gln Phe Gly Asn Val Ser Tyr Gly Ile Glu Pro Leu Glu Ser Ser	
125 130 135	

- 36 -

AGT GGT TTT GAA CAC GTG ATC TAC CAA GTG GAA CCT GAG AAA GGA GGT Ser Gly Phe Glu His Val Ile Tyr Gln Val Glu Pro Glu Lys Gly Gly 140 145 150 155	481
GCA TTA CTC TAC GCC GAG AAG GAT ATC GAT TTA AGA GAC TCG CAG TAT Ala Leu Leu Tyr Ala Glu Lys Asp Ile Asp Leu Arg Asp Ser Gln Tyr 160 165 170	529
AAG ATA CGA AGT ATC AAG CCA CAG CGG ATC GTC TCT CAC TAT TTG GAA Lys Ile Arg Ser Ile Lys Pro Gln Arg Ile Val Ser His Tyr Leu Glu 175 180 185	577
ATA CAT ATT GTC GTT GAA AAG CAA ATG TTT GAG CAT ATC GGG GCT GAT Ile His Ile Val Val Glu Lys Gln Met Phe Glu His Ile Gly Ala Asp 190 195 200	625
ACA GCC ATT GTC ACT CAA AAG ATT TTC CAG TTG ATT GGA CTG GCA AAT Thr Ala Ile Val Thr Gln Lys Ile Phe Gln Leu Ile Gly Leu Ala Asn 205 210 215	673
GCT ATC TTT GCC CCC TTT AAT CTT ACA GTA ATT CTG TCT TCC CTG GAA Ala Ile Phe Ala Pro Phe Asn Leu Thr Val Ile Leu Ser Ser Leu Glu 220 225 230 235	721
TTT TGG ATG GAT GAA AAC AAA ATC TTG ACC ACA GGC GAT GCT AAC AAG Phe Trp Met Asp Glu Asn Lys Ile Leu Thr Thr Gly Asp Ala Asn Lys 240 245 250	769
TTG CTC TAC AGG TTC CTG AAG TGG AAA CAG TCG TAC CTT GTT CTG CGA Leu Leu Tyr Arg Phe Leu Lys Trp Lys Gln Ser Tyr Leu Val Leu Arg 255 260 265	817
CCA CAT GAT ATG GCG TTT TTA CTC GTC TAC AGG AAC ACT ACC GAT TAT Pro His Asp Met Ala Phe Leu Leu Val Tyr Arg Asn Thr Thr Asp Tyr 270 275 280	865
GTT GGC GCT ACC TAT CAA GGG AAG ATG TGT GAC AAG AAC TAT GCA GGA Val Gly Ala Thr Tyr Gln Gly Lys Met Cys Asp Lys Asn Tyr Ala Gly 285 290 295	913
GGA GTT GCT TTG CAC CCC AAA GCC GTA ACT CTG GAA TCA CTT GCA ATT Gly Val Ala Leu His Pro Lys Ala Val Thr Leu Glu Ser Leu Ala Ile 300 305 310 315	961
ATT TTA GTT CAG CTG CTG AGC CTC AGC ATG GGG CTA GCG TAT GAC GAC Ile Leu Val Gln Leu Leu Ser Leu Ser Met Gly Leu Ala Tyr Asp Asp 320 325 330	1009
GTG AAC AAG TGC CAG TGT GGC GTA CCT GTC TGC GTG ATG AAC CCG GAA Val Asn Lys Cys Gln Cys Gly Val Pro Val Cys Val Met Asn Pro Glu 335 340 345	1057
GCG CCT CAC TCC AGC GGT GTC CGG GCC TTC AGT AAC TGC AGC ATG GAG Ala Pro His Ser Ser Gly Val Arg Ala Phe Ser Asn Cys Ser Met Glu 350 355 360	1105

- 37 -

GAC Asp	TTT Phe	TCC Ser	AAG Lys	TTT Phe	ATC Ile	ACA Thr	AGT Ser	CAA Gln	AGC Ser	TCC Ser	CAC His	TGT Cys	CTG Leu	CAG Gln	AAC Asn	1153
365						370					375					
CAG Gln	CCA Pro	ACG Thr	CTA Leu	CAG Gln	CCA Pro	TCT Ser	TAC Tyr	AAG Lys	ATG Met	GCG Ala	GTC Val	TGT Cys	GGG Gly	AAT Asn	GGA Gly	1201
380					385					390					395	
GAG Glu	GTG Val	GAA Glu	GAA Glu	GAT Asp	GAA Glu	ATT Ile	TGC Cys	GAC Asp	TGT Cys	GGA Gly	AAG Lys	AAG Lys	GGC Gly	TGT Cys	GCA Ala	1249
				400					405					410		
GAA Glu	ATG Met	CCC Pro	CCG Pro	CCA Pro	TGC Cys	TGT Cys	AAC Asn	CCC Pro	GAC Asp	ACC Thr	TGT Cys	AAG Lys	CTG Leu	TCA Ser	GAT Asp	1297
			415					420					425			
GGC Gly	TCC Ser	GAG Glu	TGC Cys	TCC Ser	AGC Ser	GGG Gly	ATA Ile	TGC Cys	TGC Cys	AAC Asn	TCG Ser	TGC Cys	AAG Lys	CTG Leu	AAG Lys	1345
		430				435						440				
CGG Arg	AAA Lys	GGG Gly	GAG Glu	GTT Val	TGC Cys	AGG Arg	CTT Leu	GCC Ala	CAA Gln	GAT Asp	GAG Glu	TGT Cys	GAT Asp	GTC Val	ACA Thr	1393
	445					450				455						
GAG Glu	TAC Tyr	TGC Cys	AAC Asn	GGC Gly	ACA Thr	TCC Ser	GAA Glu	GTG Val	TGT Cys	GAA Glu	GAC Asp	TTC Phe	TTT Phe	GTT Val	CAA Gln	1441
460					465					470					475	
AAC Asn	GGT Gly	CAC His	CCA Pro	TGT Cys	GAC Asp	AAT Asn	CGC Arg	AAG Lys	TGG Trp	ATC Ile	TGT Cys	ATT Ile	AAC Asn	GGC Gly	ACC Thr	1489
				480					485					490		
TGT Cys	CAG Gln	AGT Ser	GGA Gly	GAA Glu	CAG Gln	CAG Gln	TGC Cys	CAG Gln	GAT Asp	CTA Leu	TTT Phe	GGC Gly	ATC Ile	GAT Asp	GCA Ala	1537
			495					500					505			
GGC Gly	TTT Phe	GGT Gly	TCA Ser	AGT Ser	GAA Glu	TGT Cys	TTC Phe	TGG Trp	GAG Glu	CTG Leu	AAT Asn	TCC Ser	AAG Lys	AGC Ser	GAC Asp	1585
		510					515					520				
ATA Ile	TCT Ser	GGG Gly	AGC Ser	TGT Cys	GGA Gly	ATC Ile	TCT Ser	GCT Ala	GGG Gly	GGA Gly	TAC Tyr	AAG Lys	GAA Glu	TGC Cys	CCA Pro	1633
	525					530					535					
CCT Pro	AAT Asn	GAC Asp	CGG Arg	ATG Met	TGT Cys	GGG Gly	AAA Lys	ATA Ile	ATA Ile	TGT Cys	AAA Lys	TAC Tyr	CAA Gln	AGT Ser	GAA Glu	1681
540					545					550					555	
AAT Asn	ATA Ile	CTA Leu	AAA Lys	TTG Leu	AGG Arg	TCT Ser	GCC Ala	ACT Thr	GTT Val	ATT Ile	TAT Tyr	GCC Ala	AAT Asn	ATA Ile	AGC Ser	1729
				560					565					570		
GGG Gly	CAT His	GTC Val	TGC Cys	GTT Val	TCC Ser	CTG Leu	GAA Glu	TAT Tyr	CCC Pro	CAA Gln	GGT Gly	CAT His	AAT Asn	GAG Glu	AGC Ser	1777
			575					580					585			

- 38 -

CAG AAG ATG TGG GTG AGA GAT GGA ACC GTC TGC GGG TCA AAT AAG GTT Gln Lys Met Trp Val Arg Asp Gly Thr Val Cys Gly Ser Asn Lys Val 590 595 600	1825
TGC CAG AAT CAA AAA TGT GTA GCA GAC ACT TTC TTG GGC TAT GAT TGC Cys Gln Asn Gln Lys Cys Val Ala Asp Thr Phe Leu Gly Tyr Asp Cys 605 610 615	1873
AAC CTG GAA AAA TGC AAC CAC CAT GGT GTA TGT AAT AAC AAG AAG AAC Asn Leu Glu Lys Cys Asn His His Gly Val Cys Asn Asn Lys Lys Asn 620 625 630 635	1921
TGC CAC TGT GAC CCC ACA TAC TTA CCT CCA GAT TGT AAA AGA ATG AAA Cys His Cys Asp Pro Thr Tyr Leu Pro Pro Asp Cys Lys Arg Met Lys 640 645 650	1969
GAT TCA TAT CCT GGC GGG AGC ATT GAT AGT GGC AAC AAG GAA AGG GCT Asp Ser Tyr Pro Gly Gly Ser Ile Asp Ser Gly Asn Lys Glu Arg Ala 655 660 665	2017
GAA CCC ATC CCT GTA CGG CCC TAC ATT GCA AGT CGT TAC CGC TCC AAG Glu Pro Ile Pro Val Arg Pro Tyr Ile Ala Ser Arg Tyr Arg Ser Lys 670 675 680	2065
TCT CCA CGG TGG CCA TTT TTC TTG ATC ATC CCT TTC TAC GTT GTG ATC Ser Pro Arg Trp Pro Phe Phe Leu Ile Ile Pro Phe Tyr Val Val Ile 685 690 695	2113
CTT GTC CTG ATT GGG ATG CTG GTA AAA GTC TAT TCC CAA AGG ATG AAA Leu Val Leu Ile Gly Met Leu Val Lys Val Tyr Ser Gln Arg Met Lys 700 705 710 715	2161
TGG AGA ATG GAT GAC TTC TCA AGC GAA GAG CAA TTT GAA AGT GAA AGT Trp Arg Met Asp Asp Phe Ser Ser Glu Glu Gln Phe Glu Ser Glu Ser 720 725 730	2209
GAA TCC AAA GAC TAGTCTGGAC AGATTCCACA ATGTCACAAG TAATTCTCTT Glu Ser Lys Asp 735	2261
CAGTGGACAG AAAAAAAGT GGAAAAGAAA AGCCTATGCA TTATCTTGCC TGAAAGTCAA	2321
GCCTGCATAT CGTGGTCTCC ATCAGGCCAG AAATCATATC TCTCCATTAC ACATGTATGA	2381
TACATATGTG TGTATATTAT TCCATAAATG ATTTACTTGT AAGAAATGAA TGATTATGAA	2441
TTTCATATTA TACTTTGATA TTTTACCCTA TTTCTGGTAG TCGGTAGTCA TCAATTGTAT	2501
TTTCTAGTAG GTACATTATA GAAAAGGCTA TAAGAAAATA AATGTGGTAC CA	2553

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 735 amino acids
- (B) TYPE: amino acid

- 39 -

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```

Met Arg Leu Ile Leu Leu Leu Ser Gly Leu Ser Glu Leu Gly Gly
 1           5           10           15

Leu Ser Gln Ser Gln Thr Glu Gly Thr Arg Glu Lys Leu His Val Gln
          20           25           30

Val Thr Val Pro Glu Lys Ile Arg Ser Val Thr Ser Asn Gly Tyr Glu
          35           40           45

Thr Gln Val Thr Tyr Asn Leu Lys Ile Glu Gly Lys Thr Tyr Thr Leu
          50           55           60

Asp Leu Met Gln Lys Pro Phe Leu Pro Pro Asn Phe Arg Val Tyr Ser
 65           70           75           80

Tyr Asp Asn Ala Gly Ile Met Arg Ser Leu Glu Gln Lys Phe Gln Asn
          85           90           95

Ile Cys Tyr Phe Gln Gly Tyr Ile Glu Gly Tyr Pro Asn Ser Met Val
          100          105          110

Ile Val Ser Thr Cys Thr Gly Leu Arg Gly Phe Leu Gln Phe Gly Asn
          115          120          125

Val Ser Tyr Gly Ile Glu Pro Leu Glu Ser Ser Ser Gly Phe Glu His
          130          135          140

Val Ile Tyr Gln Val Glu Pro Glu Lys Gly Gly Ala Leu Leu Tyr Ala
          145          150          155          160

Glu Lys Asp Ile Asp Leu Arg Asp Ser Gln Tyr Lys Ile Arg Ser Ile
          165          170          175

Lys Pro Gln Arg Ile Val Ser His Tyr Leu Glu Ile His Ile Val Val
          180          185          190

Glu Lys Gln Met Phe Glu His Ile Gly Ala Asp Thr Ala Ile Val Thr
          195          200          205

Gln Lys Ile Phe Gln Leu Ile Gly Leu Ala Asn Ala Ile Phe Ala Pro
          210          215          220

Phe Asn Leu Thr Val Ile Leu Ser Ser Leu Glu Phe Trp Met Asp Glu
          225          230          235          240

Asn Lys Ile Leu Thr Thr Gly Asp Ala Asn Lys Leu Leu Tyr Arg Phe
          245          250          255

Leu Lys Trp Lys Gln Ser Tyr Leu Val Leu Arg Pro His Asp Met Ala
          260          265          270

```

- 40 -

Phe Leu Leu Val Tyr Arg Asn Thr Thr Asp Tyr Val Gly Ala Thr Tyr
 275 280 285
 Gln Gly Lys Met Cys Asp Lys Asn Tyr Ala Gly Gly Val Ala Leu His
 290 295 300
 Pro Lys Ala Val Thr Leu Glu Ser Leu Ala Ile Ile Leu Val Gln Leu
 305 310 315 320
 Leu Ser Leu Ser Met Gly Leu Ala Tyr Asp Asp Val Asn Lys Cys Gln
 325 330 335
 Cys Gly Val Pro Val Cys Val Met Asn Pro Glu Ala Pro His Ser Ser
 340 345 350
 Gly Val Arg Ala Phe Ser Asn Cys Ser Met Glu Asp Phe Ser Lys Phe
 355 360 365
 Ile Thr Ser Gln Ser Ser His Cys Leu Gln Asn Gln Pro Thr Leu Gln
 370 375 380
 Pro Ser Tyr Lys Met Ala Val Cys Gly Asn Gly Glu Val Glu Glu Asp
 385 390 395 400
 Glu Ile Cys Asp Cys Gly Lys Lys Gly Cys Ala Glu Met Pro Pro Pro
 405 410 415
 Cys Cys Asn Pro Asp Thr Cys Lys Leu Ser Asp Gly Ser Glu Cys Ser
 420 425 430
 Ser Gly Ile Cys Cys Asn Ser Cys Lys Leu Lys Arg Lys Gly Glu Val
 435 440 445
 Cys Arg Leu Ala Gln Asp Glu Cys Asp Val Thr Glu Tyr Cys Asn Gly
 450 455 460
 Thr Ser Glu Val Cys Glu Asp Phe Phe Val Gln Asn Gly His Pro Cys
 465 470 475 480
 Asp Asn Arg Lys Trp Ile Cys Ile Asn Gly Thr Cys Gln Ser Gly Glu
 485 490 495
 Gln Gln Cys Gln Asp Leu Phe Gly Ile Asp Ala Gly Phe Gly Ser Ser
 500 505 510
 Glu Cys Phe Trp Glu Leu Asn Ser Lys Ser Asp Ile Ser Gly Ser Cys
 515 520 525
 Gly Ile Ser Ala Gly Gly Tyr Lys Glu Cys Pro Pro Asn Asp Arg Met
 530 535 540
 Cys Gly Lys Ile Ile Cys Lys Tyr Gln Ser Glu Asn Ile Leu Lys Leu
 545 550 555 560

- 41 -

Arg Ser Ala Thr Val Ile Tyr Ala Asn Ile Ser Gly His Val Cys Val
565 570 575

Ser Leu Glu Tyr Pro Gln Gly His Asn Glu Ser Gln Lys Met Trp Val
580 585 590

Arg Asp Gly Thr Val Cys Gly Ser Asn Lys Val Cys Gln Asn Gln Lys
595 600 605

Cys Val Ala Asp Thr Phe Leu Gly Tyr Asp Cys Asn Leu Glu Lys Cys
610 615 620

Asn His His Gly Val Cys Asn Asn Lys Lys Asn Cys His Cys Asp Pro
625 630 635 640

Thr Tyr Leu Pro Pro Asp Cys Lys Arg Met Lys Asp Ser Tyr Pro Gly
645 650 655

Gly Ser Ile Asp Ser Gly Asn Lys Glu Arg Ala Glu Pro Ile Pro Val
660 665 670

Arg Pro Tyr Ile Ala Ser Arg Tyr Arg Ser Lys Ser Pro Arg Trp Pro
675 680 685

Phe Phe Leu Ile Ile Pro Phe Tyr Val Val Ile Leu Val Leu Ile Gly
690 695 700

Met Leu Val Lys Val Tyr Ser Gln Arg Met Lys Trp Arg Met Asp Asp
705 710 715 720

Phe Ser Ser Glu Glu Gln Phe Glu Ser Glu Ser Glu Ser Lys Asp
725 730 735

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2650 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 72..2273

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CATCTCGCAC TTCCAAC TGC CCTGTAACCA CCAACTGCCC TTATTCCGGC TGGGACCCAG 60

GACTTCAAGC C ATG TGG GTC TTG TTT CTG CTC AGC GGG CTC GGC GGG CTG 110

Met Trp Val Leu Phe Leu Leu Ser Gly Leu Gly Gly Leu 740 745

- 42 -

CGG	ATG	GAC	AGT	AAT	TTT	GAT	AGT	TTA	CCT	GTG	CAA	ATT	ACA	GTT	CCG	158
Arg	Met	Asp	Ser	Asn	Phe	Asp	Ser	Leu	Pro	Val	Gln	Ile	Thr	Val	Pro	
750						755					760					
GAG	AAA	ATA	CGG	TCA	ATA	ATA	AAG	GAA	GGA	ATT	GAA	TCG	CAG	GCA	TCC	206
Glu	Lys	Ile	Arg	Ser	Ile	Ile	Lys	Glu	Gly	Ile	Glu	Ser	Gln	Ala	Ser	
765					770					775					780	
TAC	AAA	ATT	GTA	ATT	GAA	GGG	AAA	CCA	TAT	ACT	GTG	AAT	TTA	ATG	CAA	254
Tyr	Lys	Ile	Val	Ile	Glu	Gly	Lys	Pro	Tyr	Thr	Val	Asn	Leu	Met	Gln	
				785					790					795		
AAA	AAC	TTT	TTA	CCC	CAT	AAT	TTT	AGA	GTT	TAC	AGT	TAT	AGT	GGC	ACA	302
Lys	Asn	Phe	Leu	Pro	His	Asn	Phe	Arg	Val	Tyr	Ser	Tyr	Ser	Gly	Thr	
			800					805					810			
GGA	ATT	ATG	AAA	CCA	CTT	GAC	CAA	GAT	TTT	CAG	AAT	TTC	TGC	CAC	TAC	350
Gly	Ile	Met	Lys	Pro	Leu	Asp	Gln	Asp	Phe	Gln	Asn	Phe	Cys	His	Tyr	
		815					820					825				
CAA	GGG	TAT	ATT	GAA	GGT	TAT	CCA	AAA	TCT	GTG	GTG	ATG	GTT	AGC	ACA	398
Gln	Gly	Tyr	Ile	Glu	Gly	Tyr	Pro	Lys	Ser	Val	Val	Met	Val	Ser	Thr	
	830					835					840					
TGT	ACT	GGA	CTC	AGG	GGC	GTA	CTA	CAG	TTT	GAA	AAT	GTT	AGT	TAT	GGA	446
Cys	Thr	Gly	Leu	Arg	Gly	Val	Leu	Gln	Phe	Glu	Asn	Val	Ser	Tyr	Gly	
845					850					855					860	
ATA	GAA	CCC	CTG	GAG	TCT	TCA	GTT	GGC	TTT	GAA	CAT	GTA	ATT	TAC	CAA	494
Ile	Glu	Pro	Leu	Glu	Ser	Ser	Val	Gly	Phe	Glu	His	Val	Ile	Tyr	Gln	
			865					870						875		
GTA	AAA	CAT	AAG	AAA	GCA	GAT	GTT	TCC	TTA	TAT	AAT	GAG	AAG	GAT	ATT	542
Val	Lys	His	Lys	Lys	Ala	Asp	Val	Ser	Leu	Tyr	Asn	Glu	Lys	Asp	Ile	
			880					885					890			
GAA	TCA	AGA	GAT	CTG	TCC	TTT	AAA	TTA	CAA	AGC	GCA	GAG	CCA	CAG	CAA	590
Glu	Ser	Arg	Asp	Leu	Ser	Phe	Lys	Leu	Gln	Ser	Ala	Glu	Pro	Gln	Gln	
		895					900					905				
GAT	TTT	GCA	AAG	TAT	ATA	GAA	ATG	CAT	GTT	ATA	GTT	GAA	AAA	CAA	TTG	638
Asp	Phe	Ala	Lys	Tyr	Ile	Glu	Met	His	Val	Ile	Val	Glu	Lys	Gln	Leu	
	910					915					920					
TAT	AAT	CAT	ATG	GGG	TCT	GAT	ACA	ACT	GTT	GTC	GCT	CAA	AAA	GTT	TTC	686
Tyr	Asn	His	Met	Gly	Ser	Asp	Thr	Thr	Val	Val	Ala	Gln	Lys	Val	Phe	
925					930					935					940	
CAG	TTG	ATT	GGA	TTG	ACG	AAT	GCT	ATT	TTT	GTT	TCA	TTT	AAT	ATT	ACA	734
Gln	Leu	Ile	Gly	Leu	Thr	Asn	Ala	Ile	Phe	Val	Ser	Phe	Asn	Ile	Thr	
			945						950					955		
ATT	ATT	CTG	TCT	TCA	TTG	GAG	CTT	TGG	ATA	GAT	GAA	AAT	AAA	ATT	GCA	782
Ile	Ile	Leu	Ser	Ser	Leu	Glu	Leu	Trp	Ile	Asp	Glu	Asn	Lys	Ile	Ala	
			960					965						970		

- 43 -

ACC ACT GGA GAA GCT AAT GAG TTA TTA CAC ACA TTT TTA AGA TGG AAA	830
Thr Thr Gly Glu Ala Asn Glu Leu Leu His Thr Phe Leu Arg Trp Lys	
975 980 985	
ACA TCT TAT CTT GTT TTA CGT CCT CAT GAT GTG GCA TTT TTA CTT GTT	878
Thr Ser Tyr Leu Val Leu Arg Pro His Asp Val Ala Phe Leu Leu Val	
990 995 1000	
TAC AGA GAA AAG TCA AAT TAT GTT GGT GCA ACC TTT CAA GGG AAG ATG	926
Tyr Arg Glu Lys Ser Asn Tyr Val Gly Ala Thr Phe Gln Gly Lys Met	
1005 1010 1015 1020	
TGT GAT GCA AAC TAT GCA GGA GGT GTT GTT CTG CAC CCC AGA ACC ATA	974
Cys Asp Ala Asn Tyr Ala Gly Gly Val Val Leu His Pro Arg Thr Ile	
1025 1030 1035	
AGT CTG GAA TCA CTT GCA GTT ATT TTA GCT CAA TTA TTG AGC CTT AGT	1022
Ser Leu Glu Ser Leu Ala Val Ile Leu Ala Gln Leu Leu Ser Leu Ser	
1040 1045 1050	
ATG GGG ATC ACT TAT GAT GAC ATT AAC AAA TGC CAG TGC TCA GGA GCT	1070
Met Gly Ile Thr Tyr Asp Asp Ile Asn Lys Cys Gln Cys Ser Gly Ala	
1055 1060 1065	
GTC TGC ATT ATG AAT CCA GAA GCA ATT CAT TTC AGT GGT GTG AAG ATC	1118
Val Cys Ile Met Asn Pro Glu Ala Ile His Phe Ser Gly Val Lys Ile	
1070 1075 1080	
TTT AGT AAC TGC AGC TTC GAA GAC TTT GCA CAT TTT ATT TCA AAG CAG	1166
Phe Ser Asn Cys Ser Phe Glu Asp Phe Ala His Phe Ile Ser Lys Gln	
1085 1090 1095 1100	
AAG TCC CAG TGT CTT CAC AAT CAG CCT CGC TTA GAT CCT TTT TTC AAA	1214
Lys Ser Gln Cys Leu His Asn Gln Pro Arg Leu Asp Pro Phe Phe Lys	
1105 1110 1115	
CAG CAA GCA GTG TGT GGT AAT GCA AAG CTG GAA GCA GGA GAG GAG TGT	1262
Gln Gln Ala Val Cys Gly Asn Ala Lys Leu Glu Ala Gly Glu Glu Cys	
1120 1125 1130	
GAC TGT GGG ACT GAA CAG GAT TGT GCC CTT ATT GGA GAA ACA TGC TGT	1310
Asp Cys Gly Thr Glu Gln Asp Cys Ala Leu Ile Gly Glu Thr Cys Cys	
1135 1140 1145	
GAT ATT GCC ACA TGT AGA TTT AAA GCC GGT TCA AAC TGT GCT GAA GGA	1358
Asp Ile Ala Thr Cys Arg Phe Lys Ala Gly Ser Asn Cys Ala Glu Gly	
1150 1155 1160	
CCA TGC TGC GAA AAC TGT CTA TTT ATG TCA AAA GAA AGA ATG TGT AGG	1406
Pro Cys Cys Glu Asn Cys Leu Phe Met Ser Lys Glu Arg Met Cys Arg	
1165 1170 1175 1180	
CCT TCC TTT GAA GAA TGC GAC CTC CCT GAA TAT TGC AAT GGA TCA TCT	1454
Pro Ser Phe Glu Glu Cys Asp Leu Pro Glu Tyr Cys Asn Gly Ser Ser	
1185 1190 1195	

- 44 -

GCA TCA TGC CCA GAA AAC CAC TAT GTT CAG ACT GGG CAT CCG TGT GGA Ala Ser Cys Pro Glu Asn His Tyr Val Gln Thr Gly His Pro Cys Gly 1200 1205 1210	1502
CTG AAT CAA TGG ATC TGT ATA GAT GGA GTT TGT ATG AGT GGG GAT AAA Leu Asn Gln Trp Ile Cys Ile Asp Gly Val Cys Met Ser Gly Asp Lys 1215 1220 1225	1550
CAA TGT ACA GAC ACA TTT GGC AAA GAA GTA GAG TTT GGC CCT TCA GAA Gln Cys Thr Asp Thr Phe Gly Lys Glu Val Glu Phe Gly Pro Ser Glu 1230 1235 1240	1598
TGT TAT TCT CAC CTT AAT TCA AAG ACT GAT GTA TCT GGA AAC TGT GGT Cys Tyr Ser His Leu Asn Ser Lys Thr Asp Val Ser Gly Asn Cys Gly 1245 1250 1255 1260	1646
ATA AGT GAT TCA GGA TAC ACA CAG TGT GAA GCT GAC AAT CTG CAG TGC Ile Ser Asp Ser Gly Tyr Thr Gln Cys Glu Ala Asp Asn Leu Gln Cys 1265 1270 1275	1694
GGA AAA TTA ATA TGT AAA TAT GTA GGT AAA TTT TTA TTA CAA ATT CCA Gly Lys Leu Ile Cys Lys Tyr Val Gly Lys Phe Leu Leu Gln Ile Pro 1280 1285 1290	1742
AGA GCC ACT ATT ATT TAT GCC AAC ATA AGT GGA CAT CTC TGC ATT GCT Arg Ala Thr Ile Ile Tyr Ala Asn Ile Ser Gly His Leu Cys Ile Ala 1295 1300 1305	1790
GTG GAA TTT GCC AGT GAT CAT GCA GAC AGC CAA AAG ATG TGG ATA AAA Val Glu Phe Ala Ser Asp His Ala Asp Ser Gln Lys Met Trp Ile Lys 1310 1315 1320	1838
GAT GGA ACT TCT TGT GGT TCA AAT AAG GTT TGC AGG AAT CAA AGA TGT Asp Gly Thr Ser Cys Gly Ser Asn Lys Val Cys Arg Asn Gln Arg Cys 1325 1330 1335 1340	1886
GTG AGT TCT TCA TAC TTG GGT TAT GAT TGT ACT ACT GAC AAA TGC AAT Val Ser Ser Ser Tyr Leu Gly Tyr Asp Cys Thr Thr Asp Lys Cys Asn 1345 1350 1355	1934
GAT AGA GGT GTA TGC AAT AAC AAA AAG CAC TGT CAC TGT AGT GCT TCA Asp Arg Gly Val Cys Asn Asn Lys Lys His Cys His Cys Ser Ala Ser 1360 1365 1370	1982
TAT TTA CCT CCA GAT TGC TCA GTT CAA TCA GAT CTA TGG CCT GGT GGG Tyr Leu Pro Pro Asp Cys Ser Val Gln Ser Asp Leu Trp Pro Gly Gly 1375 1380 1385	2030
AGT ATT GAC AGT GGC AAT TTT CCA CCT GTA GCT ATA CCA GCC AGA CTC Ser Ile Asp Ser Gly Asn Phe Pro Pro Val Ala Ile Pro Ala Arg Leu 1390 1395 1400	2078
CCT GAA AGG CGC TAC ATT GAG AAC ATT TAC CAT TCC AAA CCA ATG AGA Pro Glu Arg Arg Tyr Ile Glu Asn Ile Tyr His Ser Lys Pro Met Arg 1405 1410 1415 1420	2126

- 45 -

TGG CCA TTT TTC TTA TTC ATT CCT TTC TTT ATT ATT TTC TGT GTA CTG	2174
Trp Pro Phe Phe Leu Phe Ile Pro Phe Phe Ile Ile Phe Cys Val Leu	
1425 1430 1435	
ATT GCT ATA ATG GTG AAA GTT AAT TTC CAA AGG AAA AAA TGG AGA ACT	2222
Ile Ala Ile Met Val Lys Val Asn Phe Gln Arg Lys Lys Trp Arg Thr	
1440 1445 1450	
GAG GAC TAT TCA AGC GAT GAG CAA CCT GAA AGT GAG AGT GAA CCT AAA	2270
Glu Asp Tyr Ser Ser Asp Glu Gln Pro Glu Ser Glu Ser Glu Pro Lys	
1455 1460 1465	
GGG TAGTCTGGAC AACAGAGATG CCATGATATC ACTTCTTCTA GAGTAATTAT	2323
Gly	
CTGTGATGGA TGGACACAAA AAAATGGAAA GAAAAGAATG TACATTACCT GGTTCCTGG	2383
GATTCAAACC TGCATATTGT GATTTTAATT TGACCAGAAA ATATGATATA TATGTATAAT	2443
TTCACAGATA ATTTACTTAT TTAAAAATGC ATGATAATGA GTTTTACATT ACAAATTTCT	2503
GTTTTTTTAA AGTTATCTTA CGCTATTTCT GTTGGTTAGT AGACACTAAT TCTGTCAGTA	2563
GGGGCATGGT ATAAGGAAAT ATCATAATGT AATGAGGTGG TACTATGATT AAAAGCCACT	2623
GTTACATTTT AAAAAAAAAA AAAAAA	2650

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 734 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Trp Val Leu Phe Leu Leu Ser Gly Leu Gly Gly Leu Arg Met Asp	
1 5 10 15	
Ser Asn Phe Asp Ser Leu Pro Val Gln Ile Thr Val Pro Glu Lys Ile	
20 25 30	
Arg Ser Ile Ile Lys Glu Gly Ile Glu Ser Gln Ala Ser Tyr Lys Ile	
35 40 45	
Val Ile Glu Gly Lys Pro Tyr Thr Val Asn Leu Met Gln Lys Asn Phe	
50 55 60	
Leu Pro His Asn Phe Arg Val Tyr Ser Tyr Ser Gly Thr Gly Ile Met	
65 70 75 80	

- 46 -

Lys Pro Leu Asp Gln Asp Phe Gln Asn Phe Cys His Tyr Gln Gly Tyr
 85 90 95
 Ile Glu Gly Tyr Pro Lys Ser Val Val Met Val Ser Thr Cys Thr Gly
 100 105 110
 Leu Arg Gly Val Leu Gln Phe Glu Asn Val Ser Tyr Gly Ile Glu Pro
 115 120 125
 Leu Glu Ser Ser Val Gly Phe Glu His Val Ile Tyr Gln Val Lys His
 130 135 140
 Lys Lys Ala Asp Val Ser Leu Tyr Asn Glu Lys Asp Ile Glu Ser Arg
 145 150 155 160
 Asp Leu Ser Phe Lys Leu Gln Ser Ala Glu Pro Gln Gln Asp Phe Ala
 165 170 175
 Lys Tyr Ile Glu Met His Val Ile Val Glu Lys Gln Leu Tyr Asn His
 180 185 190
 Met Gly Ser Asp Thr Thr Val Val Ala Gln Lys Val Phe Gln Leu Ile
 195 200 205
 Gly Leu Thr Asn Ala Ile Phe Val Ser Phe Asn Ile Thr Ile Ile Leu
 210 215 220
 Ser Ser Leu Glu Leu Trp Ile Asp Glu Asn Lys Ile Ala Thr Thr Gly
 225 230 235 240
 Glu Ala Asn Glu Leu Leu His Thr Phe Leu Arg Trp Lys Thr Ser Tyr
 245 250 255
 Leu Val Leu Arg Pro His Asp Val Ala Phe Leu Leu Val Tyr Arg Glu
 260 265 270
 Lys Ser Asn Tyr Val Gly Ala Thr Phe Gln Gly Lys Met Cys Asp Ala
 275 280 285
 Asn Tyr Ala Gly Gly Val Val Leu His Pro Arg Thr Ile Ser Leu Glu
 290 295 300
 Ser Leu Ala Val Ile Leu Ala Gln Leu Leu Ser Leu Ser Met Gly Ile
 305 310 315 320
 Thr Tyr Asp Asp Ile Asn Lys Cys Gln Cys Ser Gly Ala Val Cys Ile
 325 330 335
 Met Asn Pro Glu Ala Ile His Phe Ser Gly Val Lys Ile Phe Ser Asn
 340 345 350
 Cys Ser Phe Glu Asp Phe Ala His Phe Ile Ser Lys Gln Lys Ser Gln
 355 360 365
 Cys Leu His Asn Gln Pro Arg Leu Asp Pro Phe Phe Lys Gln Gln Ala
 370 375 380

- 47 -

Val Cys Gly Asn Ala Lys Leu Glu Ala Gly Glu Glu Cys Asp Cys Gly
 385 390 395 400
 Thr Glu Gln Asp Cys Ala Leu Ile Gly Glu Thr Cys Cys Asp Ile Ala
 405 410 415
 Thr Cys Arg Phe Lys Ala Gly Ser Asn Cys Ala Glu Gly Pro Cys Cys
 420 425 430
 Glu Asn Cys Leu Phe Met Ser Lys Glu Arg Met Cys Arg Pro Ser Phe
 435 440 445
 Glu Glu Cys Asp Leu Pro Glu Tyr Cys Asn Gly Ser Ser Ala Ser Cys
 450 455 460
 Pro Glu Asn His Tyr Val Gln Thr Gly His Pro Cys Gly Leu Asn Gln
 465 470 475 480
 Trp Ile Cys Ile Asp Gly Val Cys Met Ser Gly Asp Lys Gln Cys Thr
 485 490 495
 Asp Thr Phe Gly Lys Glu Val Glu Phe Gly Pro Ser Glu Cys Tyr Ser
 500 505 510
 His Leu Asn Ser Lys Thr Asp Val Ser Gly Asn Cys Gly Ile Ser Asp
 515 520 525
 Ser Gly Tyr Thr Gln Cys Glu Ala Asp Asn Leu Gln Cys Gly Lys Leu
 530 535 540
 Ile Cys Lys Tyr Val Gly Lys Phe Leu Leu Gln Ile Pro Arg Ala Thr
 545 550 555 560
 Ile Ile Tyr Ala Asn Ile Ser Gly His Leu Cys Ile Ala Val Glu Phe
 565 570 575
 Ala Ser Asp His Ala Asp Ser Gln Lys Met Trp Ile Lys Asp Gly Thr
 580 585 590
 Ser Cys Gly Ser Asn Lys Val Cys Arg Asn Gln Arg Cys Val Ser Ser
 595 600 605
 Ser Tyr Leu Gly Tyr Asp Cys Thr Thr Asp Lys Cys Asn Asp Arg Gly
 610 615 620
 Val Cys Asn Asn Lys Lys His Cys His Cys Ser Ala Ser Tyr Leu Pro
 625 630 635 640
 Pro Asp Cys Ser Val Gln Ser Asp Leu Trp Pro Gly Gly Ser Ile Asp
 645 650 655
 Ser Gly Asn Phe Pro Pro Val Ala Ile Pro Ala Arg Leu Pro Glu Arg
 660 665 670

- 48 -

Arg	Tyr	Ile	Glu	Asn	Ile	Tyr	His	Ser	Lys	Pro	Met	Arg	Trp	Pro	Phe
	675						680					685			
Phe	Leu	Phe	Ile	Pro	Phe	Phe	Ile	Ile	Phe	Cys	Val	Leu	Ile	Ala	Ile
	690						695				700				
Met	Val	Lys	Val	Asn	Phe	Gln	Arg	Lys	Lys	Trp	Arg	Thr	Glu	Asp	Tyr
705						710				715					720
Ser	Ser	Asp	Glu	Gln	Pro	Glu	Ser	Glu	Ser	Glu	Pro	Lys	Gly		
				725					730						

- 49 -

WHAT IS CLAIMED IS:

1. A sperm protein in substantially pure form selected from a human PH30 beta chain protein, a mouse PH30 beta chain protein or an amino acid sequence substantially homologous to either the human or mouse PH30 beta chain protein.
2. The sperm protein of Claim 1, having an integrin binding sequence which is not TDE.
3. The sperm protein of Claim 2, wherein the integrin binding sequence is selected from FEE or QDE.
4. The sperm protein of Claim 1 which is the human PH30 beta chain protein.
5. The sperm protein of Claim 4, having an integrin binding sequence which is FEE.
6. A DNA sequence which encodes the sperm protein of Claim 1 or a portion of the sperm protein sufficient to constitute at least one epitope.
7. The DNA sequence of Claim 6, wherein the epitope is on the native protein.
8. The DNA sequence of Claim 6 which encodes all or a portion of human PH30 beta chain protein.
9. The DNA sequence of Claim 8, wherein the DNA encoding all or a portion of the human PH30 beta protein is

- 50 -

characterized by the ability to hybridize, under standard conditions, to the DNA sequence shown in SEQ ID NO: 1.

- 5 10. A contraceptive composition comprising a
therapeutically effective amount of the protein of Claim 1, or a
polypeptide having the substantially same amino acid sequence as a
segment of the protein provided that the polypeptide is sufficient to
constitute at least one epitope, and a pharmaceutically acceptable
carrier.
- 10 11. The contraceptive composition of Claim 10, wherein
the epitope is on the native protein.
- 15 12. The contraceptive composition of Claim 10, wherein
the protein is the human PH30 beta chain protein.
- 20 13. The contraceptive composition of Claim 10, wherein
the protein is produced by expressing the gene encoding an
immunogenic epitope of the sperm protein in a recombinant DNA
expression vector.
- 25 14. A vector comprising an inserted DNA sequence
encoding for the protein of Claim 1.
- 30 15. The vector of Claim 14, wherein the inserted DNA
sequence is characterized by the ability to hybridize, under standard
conditions, to a DNA sequence selected from the DNA sequences of
SEQ ID NO: 1 or SEQ ID NO: 3.
16. A host that is compatible with and contains the vector
of Claim 14.

- 51 -

17. A method of producing a human or mouse PH30 beta chain sperm protein, comprising the steps of culturing cells containing the DNA of Claim 6 and recovering the sperm protein from the cell culture.

5

18. The method of Claim 17, wherein the DNA encoding all or a portion of the PH30 beta chain protein is characterized by the ability to hybridize, under standard conditions, to a DNA sequence selected from the DNA sequences of SEQ ID NO: 1 or SEQ ID NO: 3.

10

19. A method of contraception in a human or mouse subject in need thereof, comprising administering to the subject an amount of the sperm protein of Claim 1 which is effective for the stimulation of antibodies which bind to the sperm protein in vivo.

15

20. The method of Claim 19, wherein the sperm protein has an integrin binding sequence which is not TDE.

21. A PH30 beta chain protein made by the process of Claim 17.

20

22. A DNA sequence as shown in Seq. ID No. 1 encoding human PH30 beta chain protein.

23. A purified and isolated DNA sequence consisting essentially of a DNA sequence encoding a polypeptide having an amino acid sequence sufficiently duplicative of that of human or mouse PH30 beta to allow the possession of the biological property of initiating sperm-egg binding or promoting sperm-egg fusion.

25

30

24. The DNA sequence of Claim 23 wherein the amino acid sequence contains an integrin binding sequence which is not TDE.

1/31
10 30 50
1 GGCCAAGATTTTCAGAATTTCTGCCACTACCAAGGTATATTGAAGGTTATCCAAAATCT 60
GlyGlnAspPheGlnAsnPheCysHisTyrGlnGlyTyrIleGluGlyTyrProLysSer
70 90 110
61 GTGGTGATGGTTAGCACATGTACTGGACTCAGGGCGTACTACAGTTTGAAAATGTTAGT 120
ValValMetValSerThrCysThrGlyLeuArgGlyValLeuGlnPheGluAsnValSer
130 150 170
121 TATGGAATAGAACCCTGGAGTCTTCAGTTGGCTTTGAACATGTAATTTACCAAGTAAAA 180
TyrGlyIleGluProLeuGluSerSerValGlyPheGluHisValIleTyrGlnValLys
190 210 230
181 CATAAGAAAGCAGATGTTTCCTTATATAATGAGAAGGATATTGAATCAAGAGATCTGTCC 240
HisLysLysAlaAspValSerLeuTyrAsnGluLysAspIleGluSerArgAspLeuSer
250 270 290
241 TTAAATTACAAAGCGCAGGCCACAGCAAGATTTTGCAAAGTATATAGAAATGCATGTT 300
PheLysLeuGlnSerAlaGluProGlnGlnAspPheAlaLysTyrIleGluMetHisVal
310 330 350
301 ATAGTTGAAAAACAATTGTATAATCATATGGGGTCTGATACAACTGTTGTGCTCAAAAA 360
IleValGluLysGlnLeuTyrAsnHisMetGlySerAspThrThrValValAlaGlnLys

FIG.1A

2/31

370 390 410

361 GTTTCCAGTTGATTGGATTGACGAATGCTATTTTGTTCATTTAATATTACAATTATT 420
ValPheGlnLeuIleGlyLeuThrAsnAlaIlePheValSerPheAsnIleThrIleIle

430 450 470

421 CTGCTTCATTGGAGCTTTGGATAGATGAAAATAAAATTGCAACCACTGGAGAAGCTAAT 480
LeuSerSerLeuGluLeuTrpIleAspGluAsnLysIleAlaThrThrGlyGluAlaAsn

490 510 530

481 GAGTTATTACACACATTTTAAGATGGAACATCTTATCTTGTTTTACGTCCTCATGAT 540
GluLeuLeuHisThrPheLeuArgTrpLysThrSerTyrLeuValLeuArgProHisAsp

550 570 590

541 GTGGCATTTTTACTTGTTTACAGAGAAAAGTCAAATTATGTTGGTCAACCTTTCAAGGG 600
ValAlaPheLeuLeuValTyrArgGluLysSerAsnTyrValGlyAlaThrPheGlnGly

610 630 650

601 AAGATGTGTGATGCAAACTATGCAGGAGGTGTTGTTCTGCACCCAGAACCATAGTCTG 660
LysMetCysAspAlaAsnTyrAlaGlyGlyValValLeuHisProArgThrIleSerLeu

670 690 710

661 GAATCACTTGCAGTTATTTTAGCTCAATTATTGAGCCTTAGTATGGGATCACTTATGAT 720
GluSerLeuAlaValIleLeuAlaGlnLeuLeuSerLeuSerMetGlyIleThrTyrAsp

FIG.1B

3 / 3 1

730 750 770
721 GACATTAACAAATGCCAGTGCTCAGGAGCTGTCTGCATTATGAATCCAGAAGCAATTCAT 780
AspIleAsnLysCysGlnCysSerGlyAlaValCysIleMetAsnProGluAlaIleHis

790 810 830
781 TTCAGTGGTGTGAAGATCTTTAGTAAGTGCAGCTTCCAAGACTTTCACATTTTATTTCA 840
PheSerGlyValLysIlePheSerAsnCysSerPheGluAspPheAlaHisPheIleSer

850 870 890
841 AAGCAGAAGTCCCAGTGCTTTCACAATCAGCCTCGCTTAGATCCTTTTTTCAAACAGCAA 900
LysGlnLysSerGlnCysLeuHisAsnGlnProArgLeuAspProPhePheLysGlnGln

910 930 950
901 GCAGTGTGTGTAATGCAAAGCTGGAAGCAGGAGAGGAGTGTGACTGTGGGACTGAACAG 960
AlaValCysGlyAsnAlaLysLeuGluAlaGlyGluGluCysAspCysGlyThrGluGln

970 990 1010
961 GATTGTGCCCTTATTGGAGAAACATGCTGTGATATTGCCACATGTAGATTTAAAGCCGGT 1020
AspCysAlaLeuIleGlyGluThrCysCysAspIleAlaThrCysArgPheLysAlaGly

1030 1050 1070
1021 TCAAAGTGTGCTGAAGGACCATGCTGCCAAAAGTGTCTATTTATGTCAAAGAAAGAATG 1080
SerAsnCysAlaGluGlyProCysCysGluAsnCysLeuPheMetSerLysGluArgMet

FIG.1C

4/31

1090	1110	1130
.		
1081	TGTAGGCCTTCCTTTGAAGAATGCGACCTCCCTGAATATTGCAATGGATCATCTGCATCA	1140
	CysArgProSerPheGluGluCysAspLeuProGluTyrCysAsnGlySerSerAlaSer	
1150	1170	1190
.		
1141	TGCCCAGAAAACCACTATGTTTCAGACTGGGCATCCGTGTGGACTGAATCAATGGATCTGT	1200
	CysProGluAsnHisTyrValGlnThrGlyHisProCysGlyLeuAsnGlnTrpIleCys	
1210	1230	1250
.		
1201	ATAGATGGAGTTTGTATGAGTGGGATAAACAATGTACAGACACATTTGGCAAAGAAGTA	1260
	IleAspGlyValCysMetSerGlyAspLysGlnCysThrAspThrPheGlyLysGluVal	
1270	1290	1310
.		
1261	GAGTTTGGCCCTTCAGAATGTTATTCTCACCTTAATTCAAAGACTGATGTATCTGGAAAC	1320
	GluPheGlyProSerGluCysTyrSerHisLeuAsnSerLysThrAspValSerGlyAsn	
1330	1350	1370
.		
1321	TGTGGTATAAGTGATTCAGGATACACACAGTGTGAAGCTGACAATCTCCAGTCCGAAAA	1380
	CysGlyIleSerAspSerGlyTyrThrGlnCysGluAlaAspAsnLeuGlnCysGlyLys	
1390	1410	1430
.		
1381	TTAATATGTAAATATGTAGCTAAATTTTATTACAAATCCAAGAGCCACTATTATTTAT	1440
	LeuIleCysLysTyrValGlyLysPheLeuLeuGlnIleProArgAlaThrIleIleTyr	

FIG.1D

5/31

1450 1470 1490

1441 GCCAACATAAGTGGACATCTCTGCATTGCTGTGGAATTTGCCAGTGATCATGCAGACAGC 1500

AlaAsnIleSerGlyHisLeuCysIleAlaValGluPheAlaSerAspHisAlaAspSer

1510 1530 1550

1501 CAAAAGATGTGGATAAAAGATGGAATTCTTGTGGTTCAAATAAGGTTTCAGGAATCAA 1560

GlnLysMetTrpIleLysAspGlyThrSerCysGlySerAsnLysValCysArgAsnGln

1570 1590 1610

1561 AGATGTGTGAGTTCTTCATACTTGGTTATGATTGTACTACTGACAAATGCAATGATAGA 1620

ArgCysValSerSerSerTyrLeuGlyTyrAspCysThrThrAspLysCysAsnAspArg

1630 1650 1670

1621 GGTGTATGCAATAACAAAAAGCACTGTCACTGTAGTGCTTCATATTTACCTCCAGATTGC 1680

GlyValCysAsnAsnLysLysHisCysHisCysSerAlaSerTyrLeuProProAspCys

1690 1710 1730

1681 TCAGTTCAATCAGATCTATGGCCTGGTGGGAGTATTGACAGTGGCAATTTCCACCTGTA 1740

SerValGlnSerAspLeuTrpProGlyGlySerIleAspSerGlyAsnPheProProVal

1750 1770 1790

1741 GCTATACCAGCCAGACTCCCTGAAAGCGCTACATTGAGAACATTTACCATTCCAAACCA 1800

AlaIleProAlaArgLeuProGluArgArgTyrIleGluAsnIleTyrHisSerLysPro

FIG.1E

6/31

1810	1830	1850
1801 ATGAGATGGCCATTTTCTTATTCATTCTTTCTTTATTATTTCTGTGTA		
1860 TGGCT 1860		
MetArgTrpProPhePheLeuPheIleProPhePheIleIlePheCysValLeuIleAla		
1870	1890	1910
1861 ATAATGGTGAAAGTTAATTTCCAAAGGAAAAATGGAGAACTGAGGACTATTCAAGCGAT 1920		
IleMetValLysValAsnPheGlnArgLysLysTrpArgThrGluAspTyrSerSerAsp		
1930	1950	1970
1921 GACCAACCTGAAAGTGAGAGTGAACCTAAAGGGTAGTCTGGACAACAGAGATGCCATGAT 1980		
GluGlnProGluSerGluSerGluProLysGly		
1990	2010	2030
1981 ATCACTTCTTCTAGAGTAATTATCTGTGATGGATGGACACAAAAAATGGAAAGAAAAGA 2040		
2050	2070	2090
2041 ATGTACATTACCTGGTTTCTGGGATTCAAACCTGCATATTGTGATTTTAATTTGACCAG 2100		
2110	2130	2150
2101 AAAATATGATATATATGTATAATTTACAGATAATTTACTTATTTAAAAATGCATGATAA 2160		

FIG.1F

7/31

2170 2190 2210
2161 TGAGTTTACATTACAAATTTCTGTTTTTTAAAGTTATCTTACGCTATTTCTGTTGGTT 2220

2230 2250 2270
2221 AGTAGACACTAATTCTGTCAGTAGGGGCATGGTATAAGGAAATATCATAATGTAATGAGG 2280

2290 2310 2330
2281 TGGTACTATGATTAAGCCACTGTTACATTTCAAAAAAAAAAAAAAAAAA 2330

FIG.1G

8/31

10 30 50

1 GGCACGAGCGATTATGTTGGCGCTACCTATCAAGGGAAGATGTGTGACAAGAACTATGCA 60
GlyThrSerAspTyrValGlyAlaThrTyrGlnGlyLysMetCysAspLysAsnTyrAla

70 90 110

61 GGAGGAGTTGCTTTGCACCCCAAAGCCGTAACCTCTGGAATCACTTGCAATTATTTAGTT 120
GlyGlyValAlaLeuHisProLysAlaValThrLeuGluSerLeuAlaIleIleLeuVal

130 150 170

121 CAGCTGCTGAGCCTCAGCATGGGGCTAGCGTATGACGACGTGAACAAGTGCCAGTGTGCC 180
GlnLeuLeuSerLeuSerMetGlyLeuAlaTyrAspAspValAsnLysCysGlnCysGly

190 210 230

181 GTACCTGTCTGCGTGATGAACCCGGAAGCGCCTCACTCCAGCGGTGTCCGGGCCTTCAGT 240
ValProValCysValMetAsnProGluAlaProHisSerSerGlyValArgAlaPheSer

250 270 290

241 AACTGCAGCATGGAGGACTTTTCCAAGTTTATCACAAGTCAAAGCTCCCACTGTCTGCAG 300
AsnCysSerMetGluAspPheSerLysPheIleThrSerGlnSerSerHisCysLeuGln

310 330 350

301 AACCAGCCAACGCTACAGCCATCTTACAAGATGGCGGTCTGTGGAATGGAGAGGTGGAA 360
AsnGlnProThrLeuGlnProSerTyrLysMetAlaValCysGlyAsnGlyGluValGlu

FIG.2A

9/31
370 390 410
361 GAAGATGAAATTTGCCACTGTGGAAAGAAGGGCTGTGCAGAAATGCCCCGCCATGCTGT 420
GluAspGluIleCysAspCysGlyLysLysGlyCysAlaGluMetProProProCysCys
430 450 470
421 AACCCCGACACCTGTAAGCTGTCAGATGGCTCCGAGTGCTCCAGCGGATATGCTGCAAC 480
AsnProAspThrCysLysLeuSerAspGlySerGluCysSerSerGlyIleCysCysAsn
490 510 530
481 TCGTGCAAGCTGAAGCGGAAAGGGAGGTTTGCAGGCTGCCCAAGATGAGTGTGATGTC 540
SerCysLysLeuLysArgLysGlyGluValCysArgLeuAlaGlnAspGluCysAspVal
550 570 590
541 ACAGAGTACTGCAACGGCACATCCGAAGTGTGTGAAGACTTCTTTGTTCAAAACGGTCAC 600
ThrGluTyrCysAsnGlyThrSerGluValCysGluAspPhePheValGlnAsnGlyHis
610 630 650
601 CCATGTGACAATCGCAAGTGGATCTGTATTAACGGCACCTGTCAGAGTGGAGAACAGCAG 660
ProCysAspAsnArgLysTrpIleCysIleAsnGlyThrCysGlnSerGlyGluGlnGln
670 690 710
661 TGCCAGGATCTATTTGGCATCGATGCAGGCTTTGGTTCAAGTGAATGTTTCTGGGAGCTG 720
CysGlnAspLeuPheGlyIleAspAlaGlyPheGlySerSerGluCysPheTrpGluLeu

FIG.2B

10 / 3 1

730 750 770

721 AATTCCAAGAGCGACATATCTGGGAGCTGTGGAATCTCTGCTGGGGATACAAGGAATGC 780
AsnSerLysSerAspIleSerGlySerCysGlyIleSerAlaGlyGlyTyrLysGluCys

790 810 830

781 CCACCTAATGACCGGATGTGTGGGAAAATAATATGTAAATACCAAAGTGAAAATATACTA 840
ProProAsnAspArgMetCysGlyLysIleIleCysLysTyrGlnSerGluAsnIleLeu

850 870 890

841 AAATTGAGGTCTGCCACTGTTATTTATGCCAATATAAGCGGCATGTCTGCGTTTCCCTG 900
LysLeuArgSerAlaThrValIleTyrAlaAsnIleSerGlyHisValCysValSerLeu

910 930 950

901 GAATATCCCCAAGGTCATAATGAGAGCCAGAAGATGTGGGTGAGAGATGGAACCGTCTGC 960
GluTyrProGlnGlyHisAsnGluSerGlnLysMetTrpValArgAspGlyThrValCys

970 990 1010

961 GGGTCAAATAAGGTTGCCAGAATCAAAAATGTGTAGCAGACACTTTCTGGGCTATGAT 1020
GlySerAsnLysValCysGlnAsnGlnLysCysValAlaAspThrPheLeuGlyTyrAsp

1030 1050 1070

1021 TGCAACCTGGA AAAATGCAACCACCATGGTGTATGTAATAACAAGAAGAACTGCCACTGT 1080
CysAsnLeuGluLysCysAsnHisHisGlyValCysAsnAsnLysLysAsnCysHisCys

FIG.2C

11/31

	1090	1110	1130	
1081	GACCCACATACTTACCTCCAGATTGTAAAAGAATGAAAGATTCATATCCTGGCGGGAGC			1140
	AspProThrTyrLeuProProAspCysLysArgMetLysAspSerTyrProGlyGlySer			
	1150	1170	1190	
1141	ATTGATAGTGGCAACAAGGAAAGGGCTGAACCCATCCCTGTACGGCCCTACATTGCAAGT			1200
	IleAspSerGlyAsnLysGluArgAlaGluProIleProValArgProTyrIleAlaSer			
	1210	1230	1250	
1201	CGTTACCGCTCCAAGTCTCCACGGTGGCCATTTTCTTGATCATCCCTTTCTACGTTGTG			1260
	ArgTyrArgSerLysSerProArgTrpProPhePheLeuIleIleProPheTyrValVal			
	1270	1290	1310	
1261	ATCCTTGTCTGATTGGGATGCTGGTAAAAGTCTATTCCCAAAGGATGAAATGGAGAATG			1320
	IleLeuValLeuIleGlyMetLeuValLysValTyrSerGlnArgMetLysTrpArgMet			
	1330	1350	1370	
1321	GATGACTTCTCAAGCGAAGAGCAATTTGAAAGTGAAAGTGAATCCAAAGACTAGTCTGGA			1380
	AspAspPheSerSerGluGluGlnPheGluSerGluSerGluSerLysAsp			
	1390	1410	1430	
1381	CAGATTCCACAATGTCACAAGTAATTCTCTTCAGTGGACAGAAAAAAGTGGAAGAA			1440
	1450	1470	1490	
1441	AAGCCTATGCATTATCTTGCTGAAAGTCAAGCCTGCATATCGTGGTCTCCATCAGGCCA			1500
	1510	1530	1550	
1501	GAAATCATATCTCTCCATTACACATGTATGATACATATGTGTGTATATTATCCATAAAT			1560

FIG.2D

12/31

1570 1590 1610
1561 GATTTACTTGTAAGAAATGAATGATTATGAATTCATATTATACTTTGATATTTTACCCT 1620
1630 1650 1670
1621 ATTTCTGGTAGTCGGTAGTCATCAATTGTATTTCTAGTAGGTACATTATAGAAAAGGCT 1680
1690
1681 ATAAGAAAATAAATGTGGTACCA 1703

FIG.2E

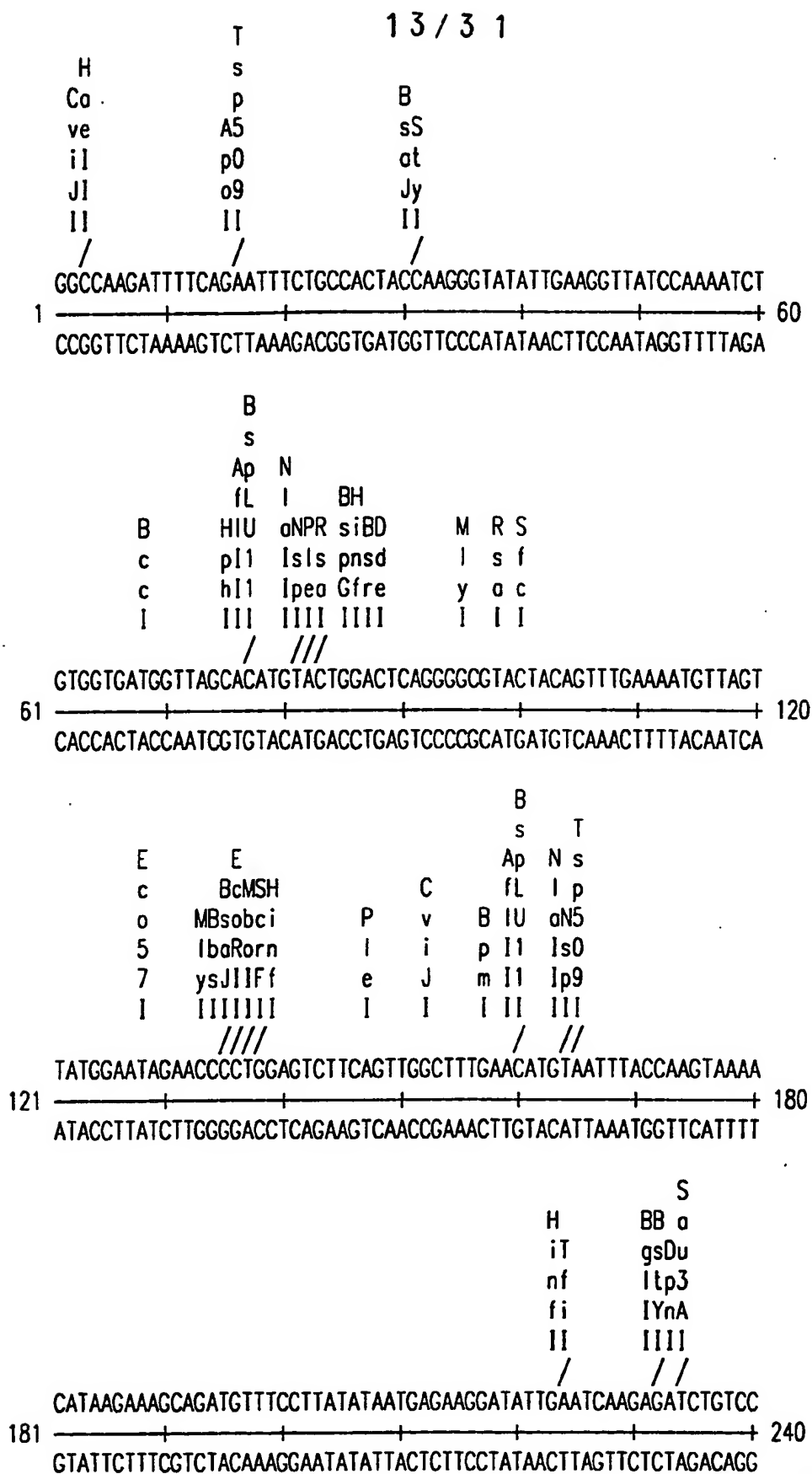


FIG.3A

14/31

T s p MD5 sr0 eo9 III	H h a I	C v i J I	C v i R I	C v i R I	N I a s l p I
---	------------------	-----------------------	-----------------------	-----------------------	---------------------------------

/

241
TTTAAATTACAAAGCGCAGAGCCACAGCAAGATTTTGCAAAGTATATAGAAATGCATGTT
300

AAATTTAATGTTTCGCGTCTCGGTGTCGTTCTAAACGTTTCATATATCTTTACGTACAA

T s p M5 u0 n9 II	N d e I	B a e I
-------------------------------------	------------------	------------------

/

301
ATAGTTGAAAAACAATTGTATAATCATATGGGGTCTGATACAACTGTTGTGCGCTCAAAAA
360

TATCAACTTTTTGTTAACATATTAGTATACCCAGACTATGTTGACAACAGCGAGTTTTT

B s r I	B s t X I	B s m I	M s e I	S s p I	T s p M 5 0 9 I	Bb bo sl II
------------------	-----------------------	------------------	------------------	------------------	--------------------------------------	----------------------

/

361
GTTTTCCAGTTGATTGGATTGACGAATGCTATTTTTGTTTCATTTAATATTACAATTATT
420

CAAAAGGTCAACTAACCTAACTGCTTACGATAAAAACAAAGTAAATTATAATGTTAATAA

C Av li uJ II	T s p 5 0 9 I	C v i R I	B s r I	C Av li uJ II
---------------------------	---------------------------------	-----------------------	------------------	---------------------------

/

421
CTGTCTTCATTGGAGCTTTGGATAGATGAAAATAAAATTGCAACCACTGGAGAAGCTAAT
480

GACAGAAGTAACCTCGAAACCTATCTACTTTTATTTTAACTGGTGACCTCTTCGATTA

FIG. 3B

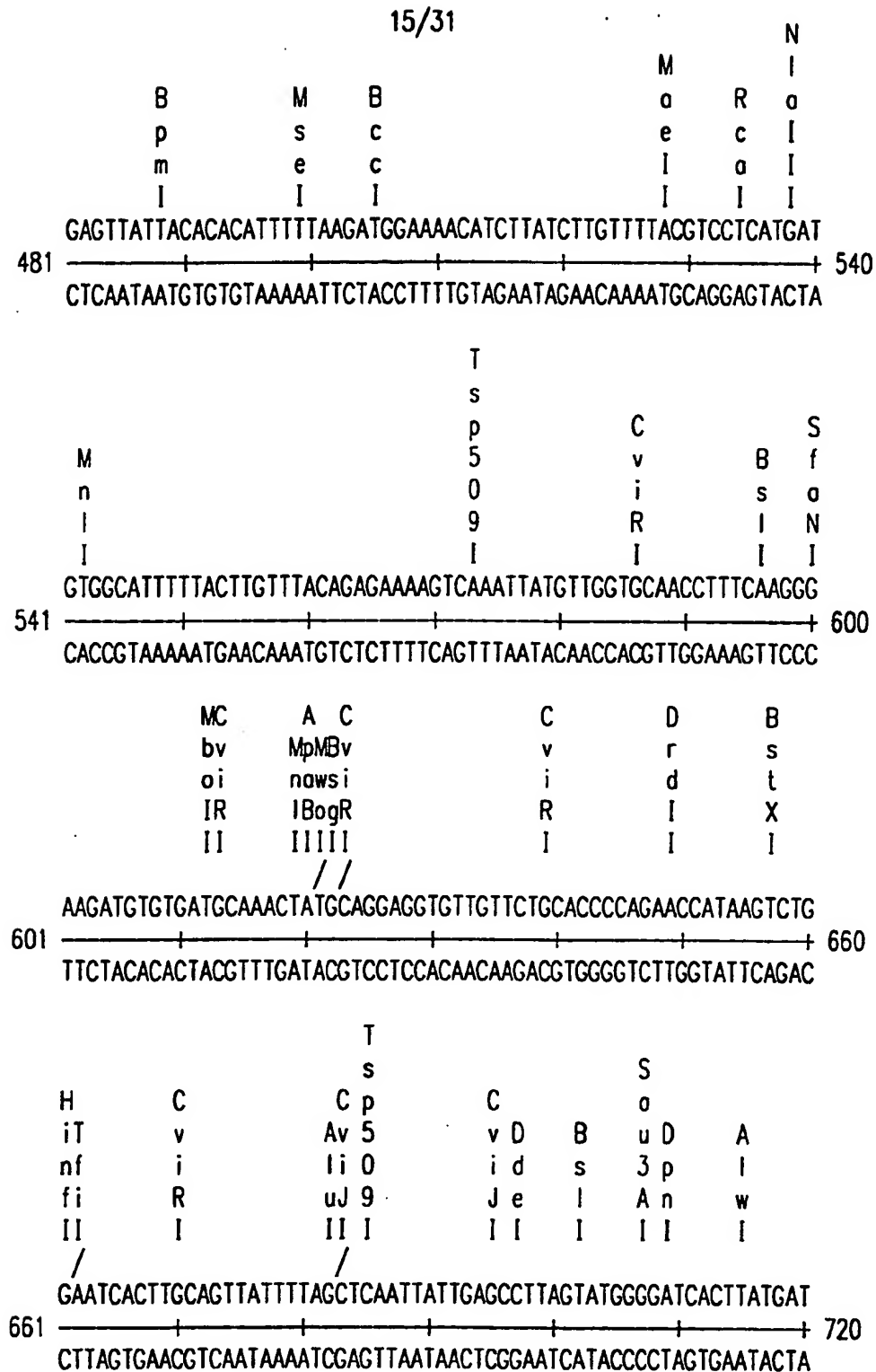


FIG.3C

16 / 3 1

		B					
		s					i
		BAp					s
		p11	AC	C	H		p
M	B	uw2D	Alv	v	iT		5X
s	s	128d	lwi	i	nf		Om
e	r	016e	uNJ	R	fi		9n
I	I	IIII	III	I	II		II
		//	//		/		/

GACATTAACAAATGCCAGTGTCTCAGGAGCTGTCTGCATTATGAATCCAGAAGCAATTCAT
 721 ————— 780
 CTGTAATTGTTTACGGTCACGAGTCTCTGACAGACGTAATACTTAGGTCTTCGTTAAGTA

	S	M	F					
	BB a	a M	Cn C			C M	B	
M	gsDu	e bB	S vuPAv	NT		B B v	b s	B
s	ltp3	I oc f	i4sli	sa		b b i	o m	c
I	IYnA	I Ig c	RHtuJ	pq		v s R	I F	g
I	IIII	I II I	IIIII	VI		I I I	I I	I
	//		//					

TTCAGTGGTGTGAAGATCTTTAGTAACTGCAGCTTCGAAGACTTTGCACATTTTATTTC
 781 ————— 840
 AAGTCACCACACTTCTAGAAATCATTGACGTCGAAGCTTCTGAAACGTGTAAAATAAAGT

			S		
	M		Bo		C
	BbB		suDM		a
	bos		t3pn		c
	slr		YAnl		8
	III		IIII		I
	/		//		

AAGCAGAAGTCCCAGTGTCTTCACAATCAGCCTCGCTTAGATCCTTTTTTCAAACAGCAA
 841 ————— 900
 TTCGTCTTCAGGTCACAGAAGTGTTAGTGGAGCGAATCTAGGAAAAAAGTTTGTCCGT

T	T				
t	t				
h	h			M T	
1	1 C	C		BR a s	
1	1 v	Av M		sl e p	
1	1 i	li n		ee 1 4	
1	1 R	uJ I		RA 1 5	
1	1 I	II I		II 1 1	
	/		/		

GCAGTGTGTGTAATGCAAAGCTGGAAGCAGGAGAGTGTGACTGTGGGACTGAACAG
 901 ————— 960
 CGTCACACACCATTACGTTTCGACCTTCGTCTCTCCTCACACTGACACCCTGACTTGTC

FIG.3D

17/31

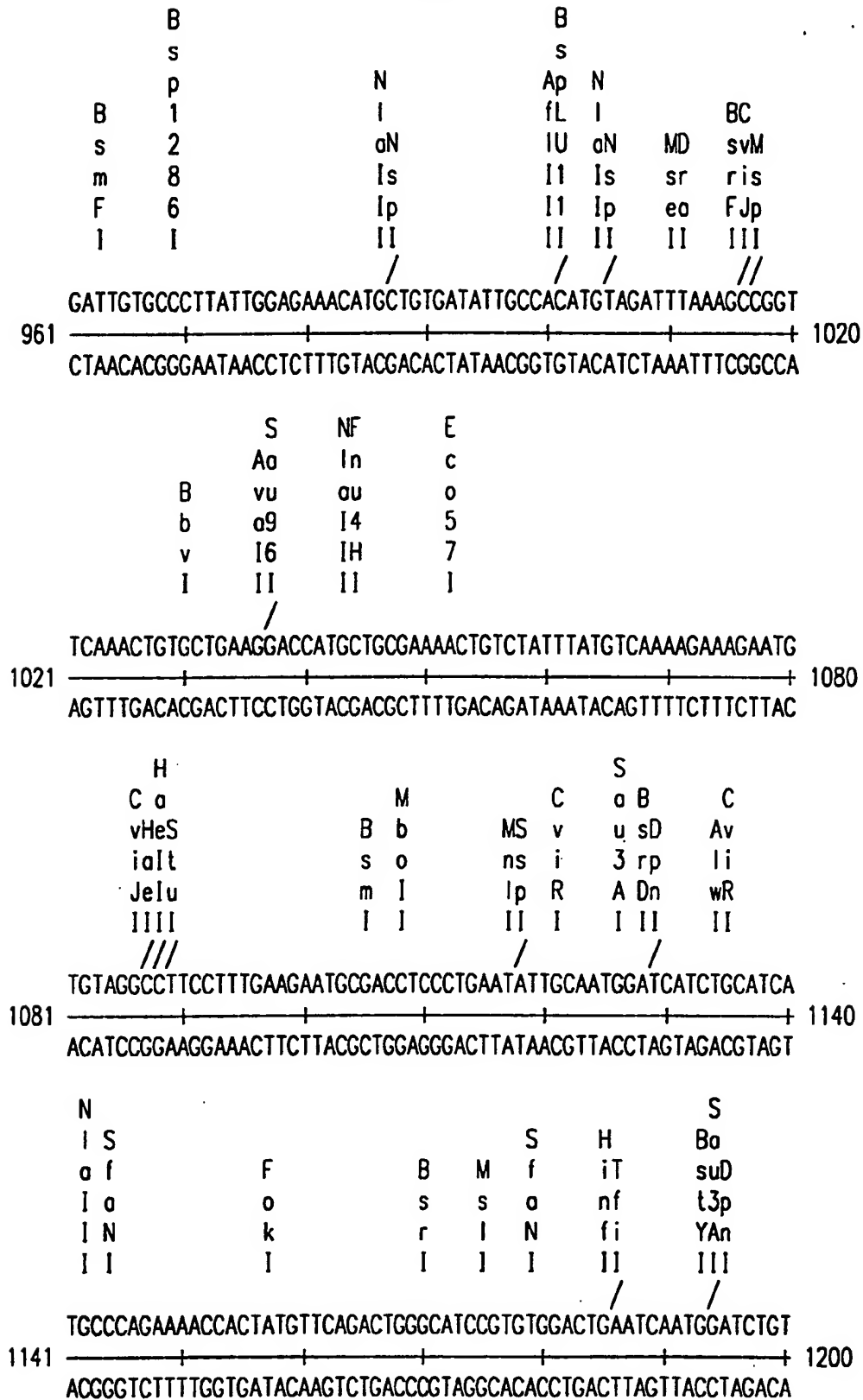


FIG.3E

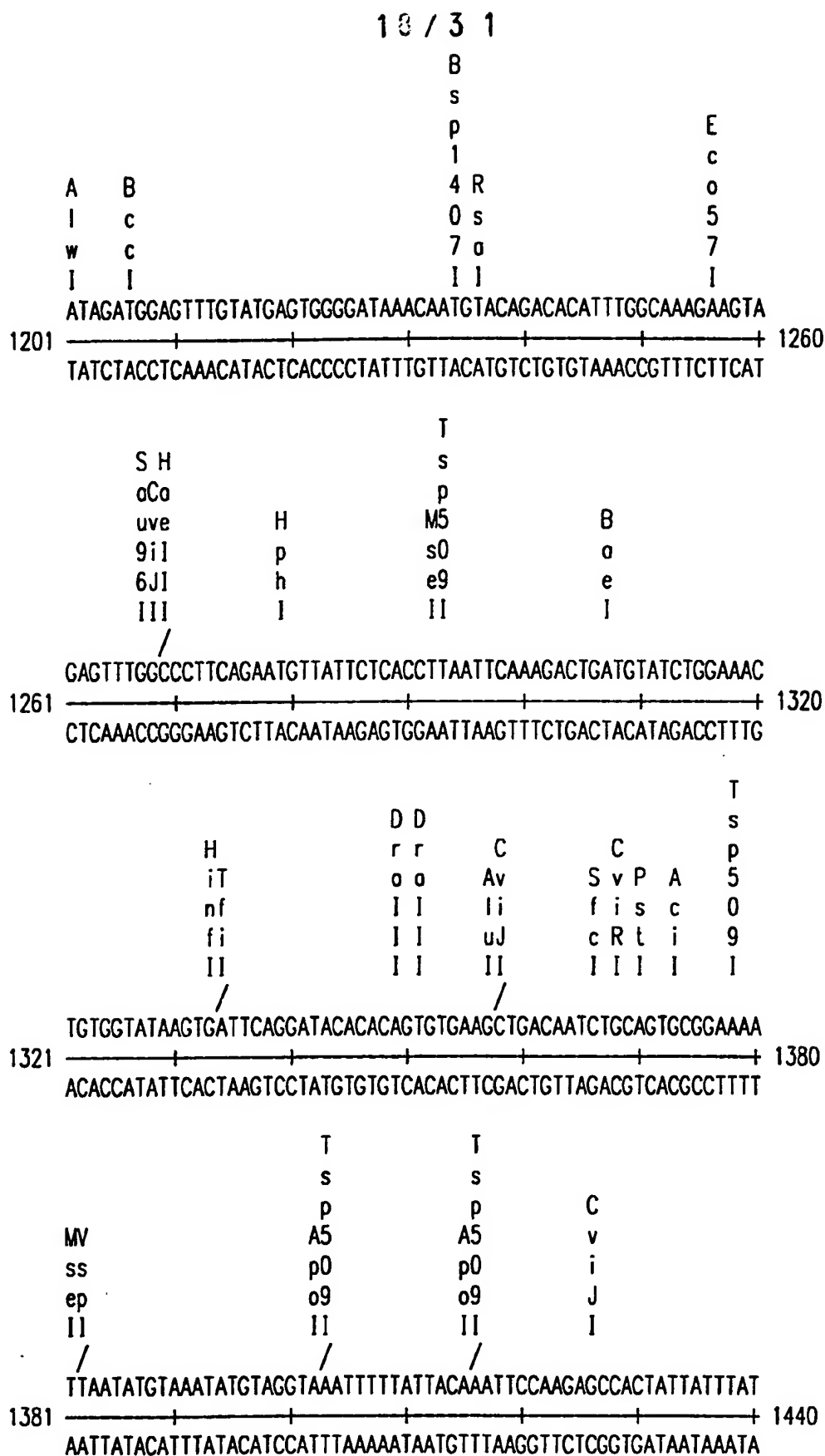


FIG.3F

19/31

T
 s
 S
 N
 C
 BC
 sv M A5 B BuD va v
 ri s p0 s c3p il i
 DR l o9 r lAn Rl J
 ll l ll l ll ll
 / / / /
 GCCAACATAAGTGGACATCTCTGCATTGCTGTGGAATTTGCCAGTGATCATGCAGACAGC
 1441 1500
 CGGTTGTATTACCTGTAGAGACGTAACGACACCTTAAACGGTCACTAGTACGTCTGTGC
 D
 C H
 B r v iT
 c d i nf
 c l R fi
 l l l ll
 /
 CAAAAGATGTGCATAAAAGATGGAACCTCTTGTGGTTCAAATAAGGTTTGCAGGAATCAA
 1501 1560
 GTTTTCTACACCTATTTTCTACCTTGAAGAACACCAAGTTTATCCAAACGTCCTTAGTT
 M
 C B
 b R v M s
 o s i n r
 l a R l D
 l l l l
 AGATGTGTGAGTTCTTCATACTTGGGTTATGATTGTACTACTGACAAATGCAATGATAGA
 1561 1620
 TCTACACACTCAAGAAGTATGAACCAATACTAACATGATGACTGTTTACGTTACTATCT
 M T
 C a s
 v e p S M B MD
 i l 4 f s p nd
 R l 5 c l m le
 l l l l l ll
 GGTGTATGCAATAACAAAAAGCACTGTCACTGTAGTGCTTCATATTTACCTCCAGATTGC
 1621 1680
 CCACATACGTTATTGTTTTTCGTGACAGTGACATCACGAAGTATAAATGGAGGTCTAACC

FIG. 3G

20/31

S E H T
 BB a Cc aS s
 gsDu voHec p 5 S
 ltp3 iRaIr 0 f
 lYnA JleIF 9 c
 llll lllll l l
 // ////
 TCAGTTC AATCAGATCTATG GCCTGGTGGGAGTATTGACAGTGGCAATTTTCCACCTGTA
 1681 |-----| 1740
 AGTCAAGTTAGTCTAGATACCGGACCACCCTCATAACTGTCACCGTTAAAGGTGGACAT

 C C H H
 Av MP v i M Ho
 li wl i n l he
 uJ oe J f y ol
 ll ll l l l ll
 /
 GCTATACCAGCCAGACTCCCTGAAAGGCGCTACATTGAGAACATTTACCATTCCAAACCA
 1741 |-----| 1800
 CGATATGGTCCGTCTGAGGGACTTTCCGCGATGTAACCTCTTGTAATGGTAAGGTTTGGT

 H
 B C a
 s BEvHeM R
 t caiaIs s
 X ceJelc a
 l llllll l
 ///
 ATGAGATGGCCATTTTCTTATTCATTCCCTTTCTTTATTATTTCTGTGTA CTGATTGCT
 1801 |-----| 1860
 TACTCTACCGGTAAAAAGAATAAGTAAAGGAAAGAAATAATAAAAAGACACATGACTAACCA

 T
 s
 p
 M5 H M D
 s0 p n d
 e9 h l e
 ll l ll ll
 ATAATGGTGAAGTAAATTTCCAAAGGAAAAAATGGAGAACTGAGGACTATTCAAGCGAT
 1861 |-----| 1920
 TATTACCACTTTCAATTAAAGGTTTCCTTTTTTACCTCTTGACTCCTGATAAGTTCGCTA

FIG. 3H

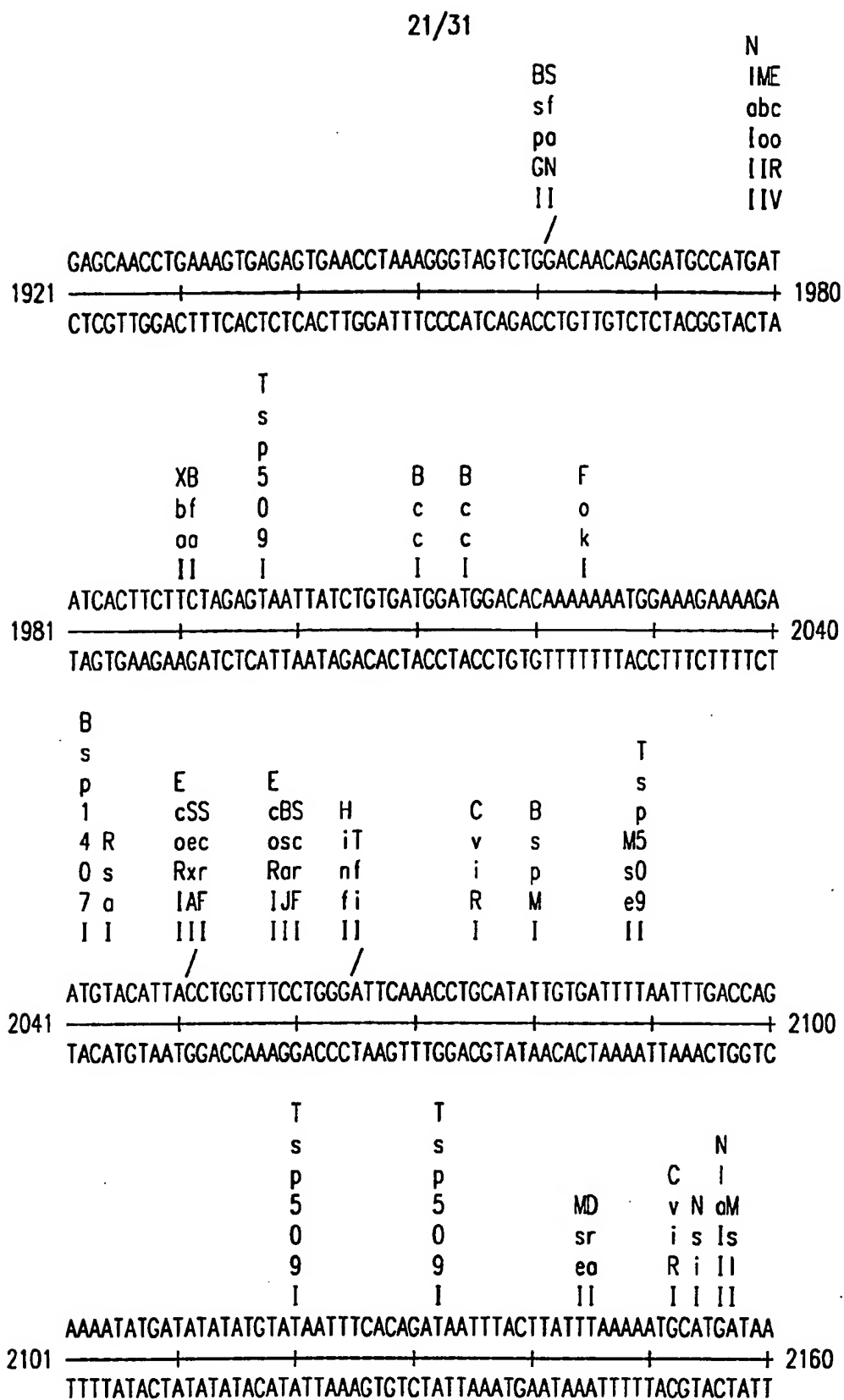
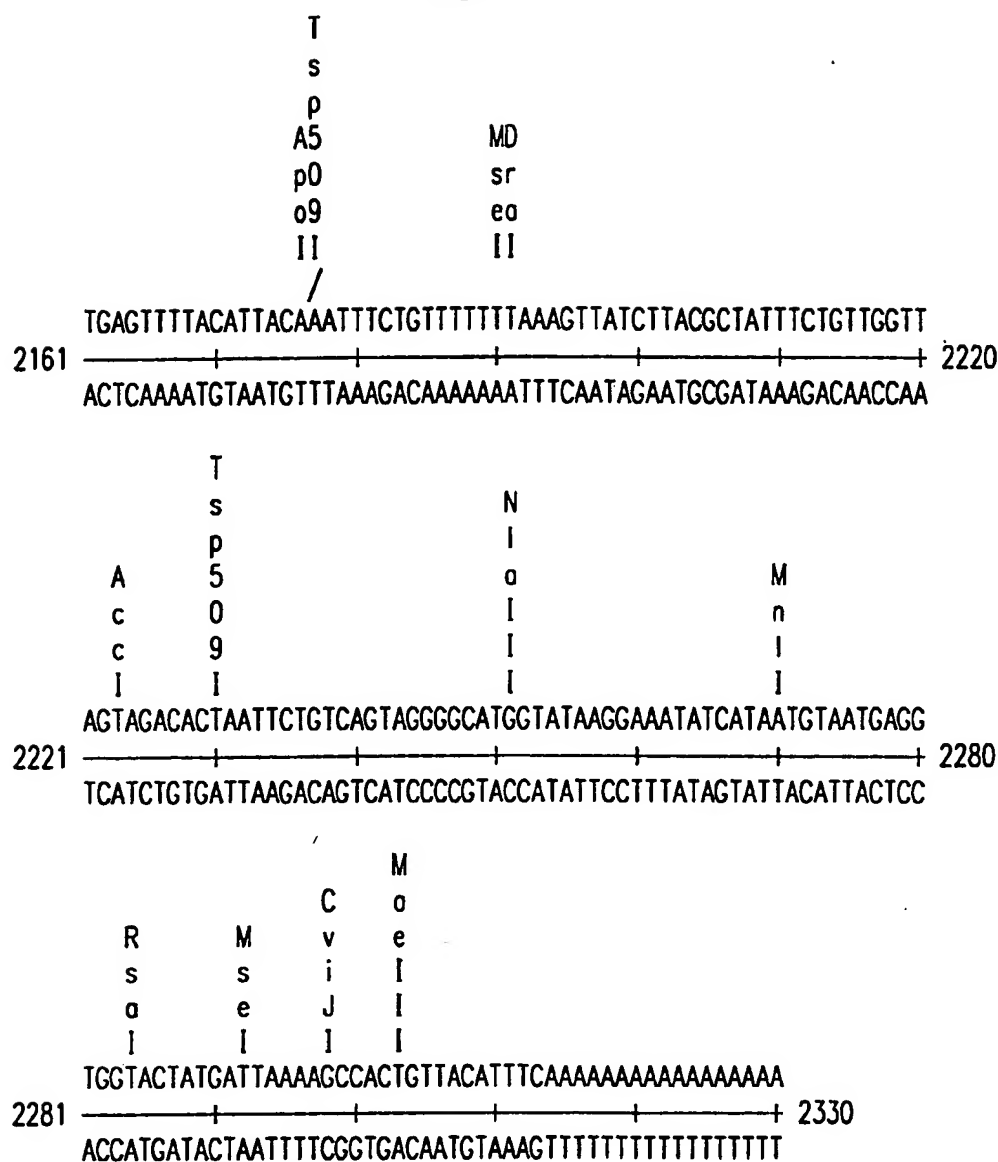


FIG.31

22 / 3 1



Enzymes that do cut:

AccI	Acil	AflIII	AluI	AlwI	Alw21I	AlwNI	ApoBI
ApoI	Avall	BaeI	BbsI	BbvI	BccI	BcgI	BclI
BfoI	BglII	BpmI	Bpu10I	BsaJI	BseRI	BsgI	BslI
BsmI	BsmFI	Bsp1286I	Bsp1407I	BspGI	BspLU11I	BspMI	BsrI
BsrDI	BsrFI	BstXI	BstYI	Cac8I	CviJI	CviRI	DdeI
DpnI	DraI	DraIII	DrdII	EaeI	Eco57I	EcoRII	EcoRV
Fnu4HI	FokI	HaeI	HaeII	HaeIII	HhaI	HinfI	HphI
MaeII	MaeIII	MboII	MlyI	MnlI	MscI	MseI	MslI
MspI	MunI	MwoI	NdeI	NlaIII	NsiI	NspI	NspV
PleI	PstI	RcoI	RleAI	RsaI	Sau96I	Sau3AI	ScrFI
SexAI	SfoNI	SfcI	SspI	StuI	StyI	TaqI	TfiI
Tsp45I	Tsp509I	Tth111II	VspI	XbaI	XmnI		

FIG.3J

23/31

Enzymes that do not cut:

AatII	AflII	AgeI	Alw44I	ApaI	AscI	AvaI	AvrII
BamHI	BanI	BanII	Bce83I	BceII	BglII	Bpu1102I	BsoI
BsaAI	BsaBI	BsaHI	BsaWI	BscGI	BsiI	BsiEI	BsiWI
BsmAI	BspEI	BsrBI	BssHII	Bst1107I	BstEII	Bsu36I	ClaI
DrdI	DsaI	EagI	Eam1105I	EorI	EciI	Eco47III	Eco105I
EcoNI	EcoO109I	EcoRI	Esp3I	FauI	FseI	FspI	GdiII
HgaI	HgiEII	HincII	HindIII	HpaI	KpnI	MluI	MmeI
NaeI	NarI	NciI	NcoI	NheI	NlaIV	NotI	NruI
NspBII	PacI	Pf11108I	PfIMI	PmeI	PmlI	PshAI	Psp5II
Psp1406I	PvuI	PvuII	RsrII	SacI	SacII	SaiI	SapI
ScaI	SfiI	SgrAI	SmaI	SpeI	SphI	SrfI	Sse8387I
SwaI	TaqII	TaqII	ThaI	Tth111I	XcmI	XhoI	

FIG.3K

B H B M T C
 s Ha s e p b M v
 i he s l 4 o n i
 l ol l l 5 l l R
 l ll l l l l l

GGCACGAGCGATTATGTTGGCGCTACCTATCAAGGGAAGATGTGTGACAAGAACTATGCA
 1 60
 CCGTGCTCGCTAATACAACCGGATGGATAGTTCCTTCTACACACTGTTCTTGATACGT

B B C M T
 s c v C a s
 e e i v e Cp
 R f R J l nf v5 B
 l l l l l fi R9 v
 l l l l l ll ll l

GGAGGAGTTGCTTTGCACCCCAAAGCCGTAACCTCTGGAATCACTTGCAATTATTTAGTT
 61 120
 CCTCCTCAACGAAACGTGGGGTTTCGGCATTGAGACCTTAGTGAACGTTAATAAAATCAA

B
 p
 F u N B N
 CnA1 sPC p l C C M
 Avu11Dpvv uD B aM vBoN a B BM
 li4wOdBui ld s In ifch e s gw
 uJHN2e11J 0e l ll Ja8e l r lo
 llllllllll ll l ll llll l ll
 // /// / / /

CAGCTGCTGAGCCTCAGCATGGGCTAGCGTATGACGACGTGAACAAGTCCAGTGTGGC
 21 180
 GTGACGACTCGGAGTCTGACCCGATCGCATACTGCTGCACTTGTTACAGGTCACACCG

R S H E N S H M
 s BMNc Ha c s aSc a
 a pscr he o AMp MNucve e
 l mpiF ol 5 cnB sc9ri l l
 l llll ll 7 ill pi6FJl l
 l llll ll l lll llllll l
 // // //

GTACCTGTCTGCGTGATGAACCGGAAGCGCCTCACTCCAGCGGTGTCGGGGCCTTCAGT
 81 240
 CATGGACAGACGCACTACTTGGGCCTTCGGGAGTGAGGTCGCCACAGGCCCGGAAGTCA

FIG. 4A

26/31

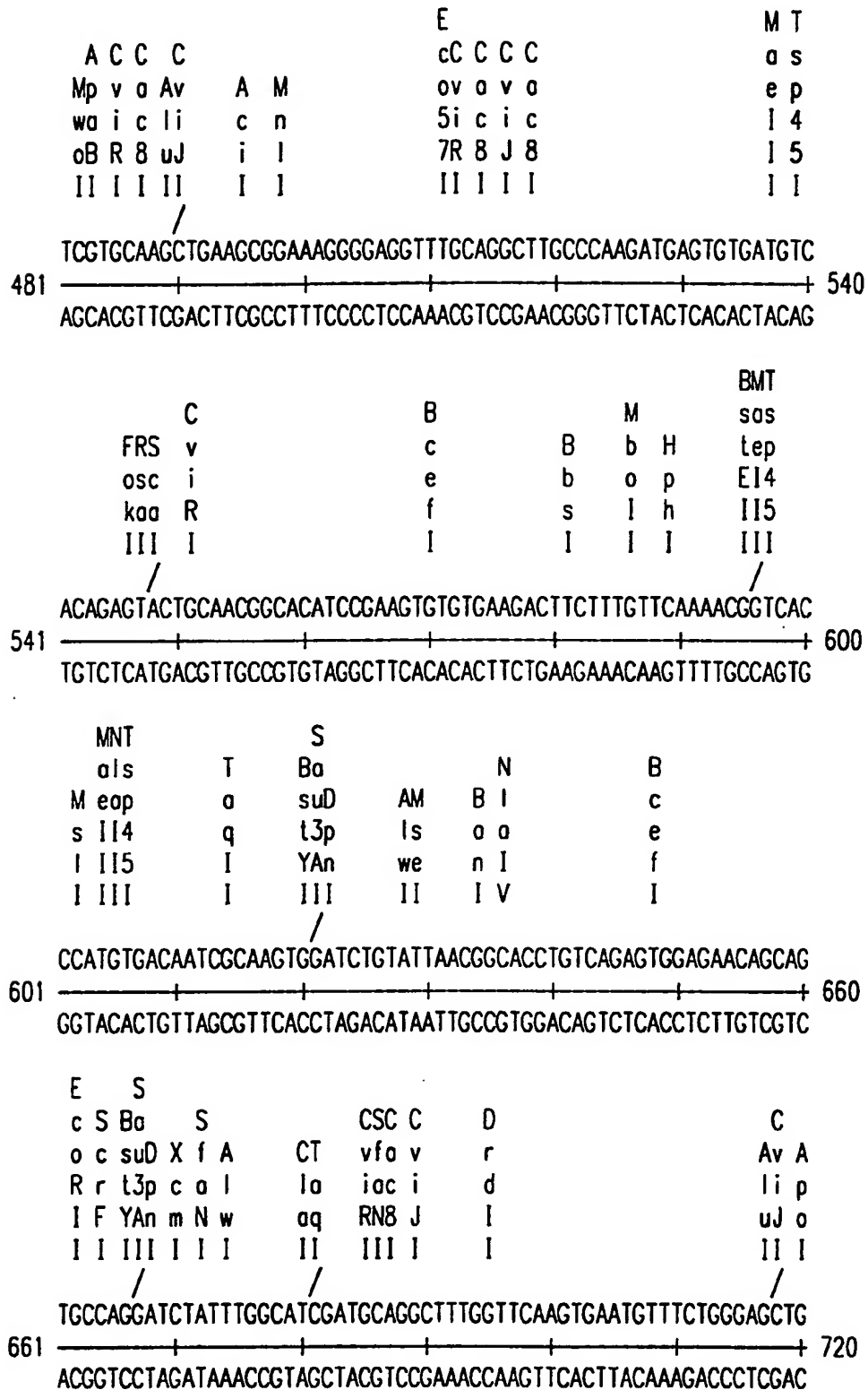


FIG.4C

T
s
Ep C H
c5 Av iT
o0 li nf
R9 uJ fi
II II II
// / /

AATTCCAAGAGCGACATATCTGGGAGCTGTGGAATCTCTGCTGGGGATACAAGGAATGC
721 780
TTAAGGTTCTCGCTGTATAGACCTCGACACCTTAGAGACGACCCCTATGTTCTTACG

R B
B I sM F M
s e as o n
m A Wp k I
I I II I I

CCACCTAATGACCGGATGTGTGGGAAAATAATATGTAATACCAAAGTGAAAATATACTA
781 840
GGTGGATTACTGGCCTACACACCTTTTATTATACATTTATGGTTTCACITTTATATGAT

T
s
p
5
0
9
I

N
C
F A a
a c c
u i 8
I I I

E
BcS
soc
aRr
JIF
III

/ /

AAATTGAGGTCTGCCACTGTTATTTATGCCAATATAAGCGGGCATGTCTCGGTTTCCCTG
841 900
TTTAACTCCAGACGGTGACAATAAATACGGTTATATTCGCCCGTACAGACGCAAAGGGAC

B RCT M N
sS lva b B HF I AB
at ei q o c p o o cs
Jy AJI I c hu I i I
II III I I IV II

/ / /

GAATATCCCCAAGGTCATAATGAGAGCCAGAAGATGTGGGTGAGAGATGGAACCGTCTGC
901 960
CTTATAGGGGTTCCAGTATTACTCTCGGTCTTCTACACCACTCTCTACCTTGGCAGACG

FIG. 4D

28/31

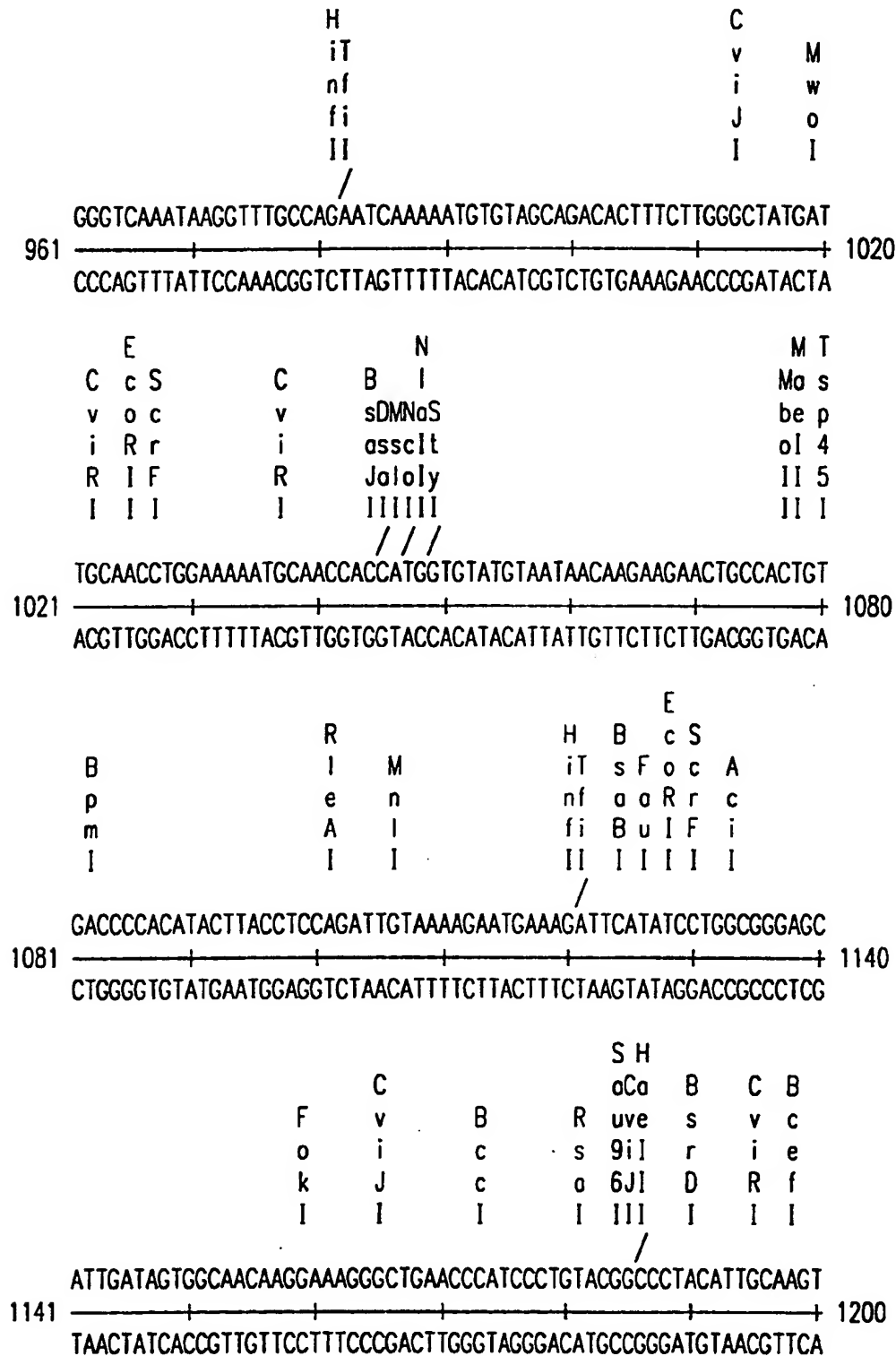


FIG.4E

29/31

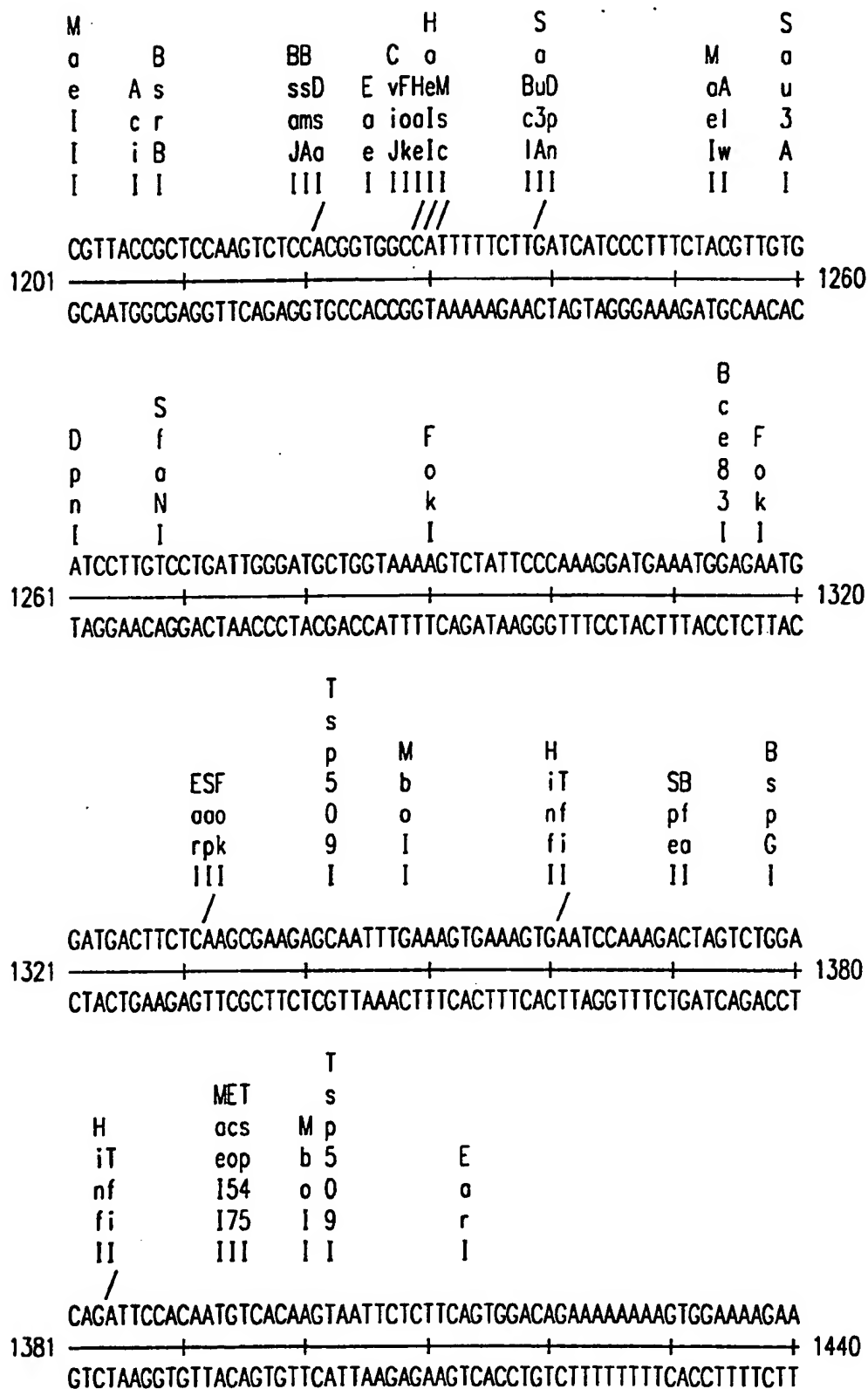


FIG.4F

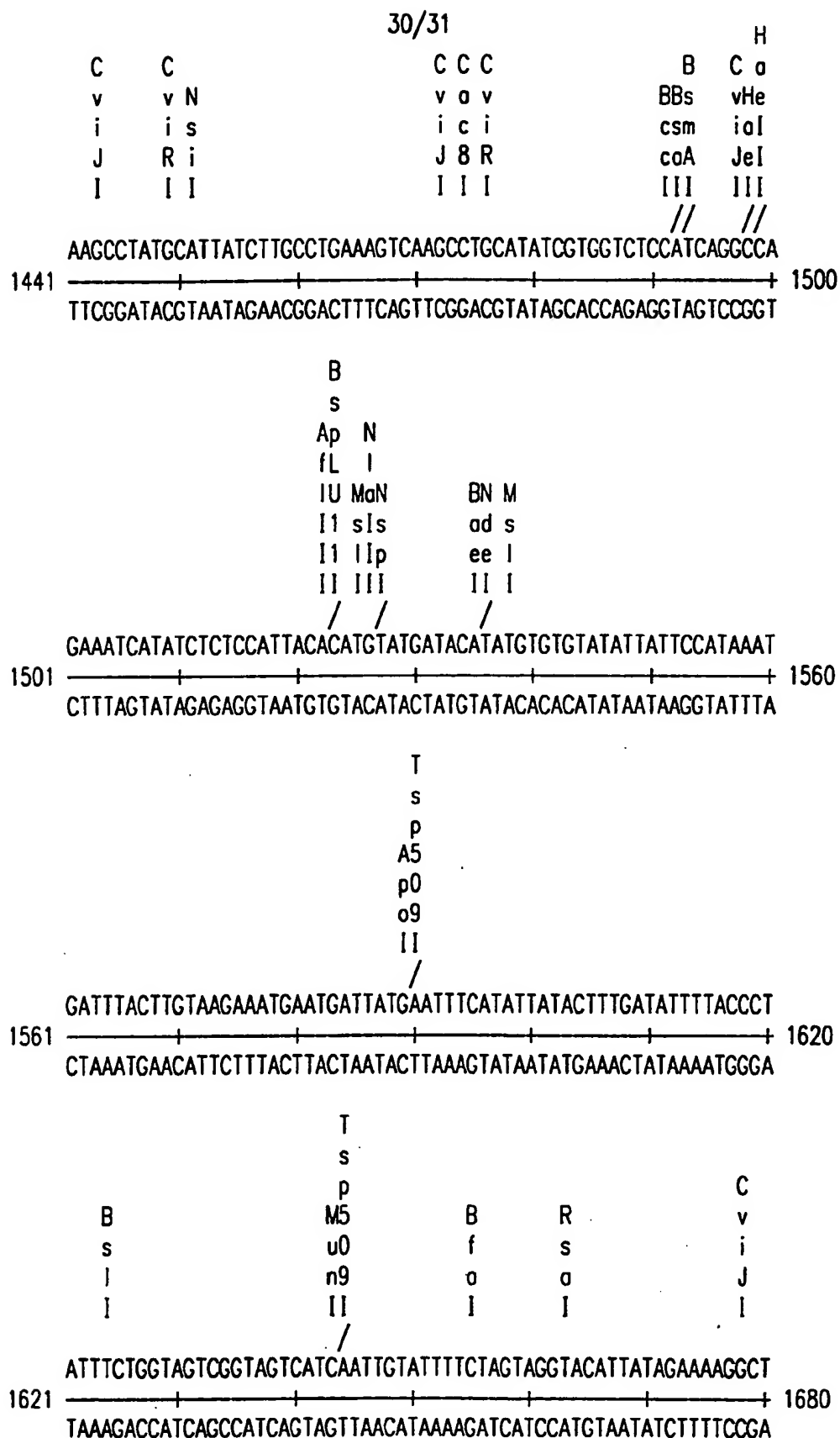
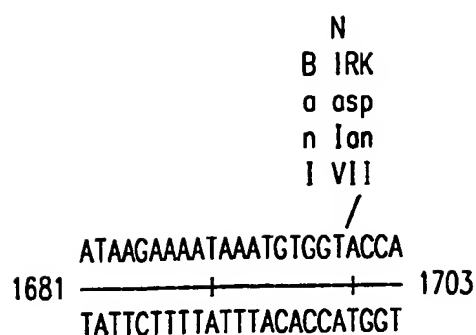


FIG.4G

31/31



Enzymes that do cut:

Acil	AflIII	AluI	AlwI	Alw21I	AlwNI	ApaBI	ApoI
BaeI	BanI	BbsI	BbvI	BccI	Bce83I	BceII	BclI
Bfal	BglI	BpmI	Bpu10I	Bpu1102I	BsaI	BsaBI	BsaJI
BsaWI	BseRI	BsgI	BsII	BsII	BsmI	BsmAI	Bsp1286I
BspGI	BspLU11I	BsrI	BsrBI	BsrDI	BstEII	BstYI	Cac8I
Clal	CviJI	CviRI	DdeI	DpnI	DrdII	DsaI	EaeI
EarI	Eco57I	EcoRI	EcoRII	FauI	Fnu4HI	FokI	HaeI
HaeII	HaeIII	HhaI	HinI	HphI	KpnI	MaeII	MaeIII
MboII	MnII	MscI	MseI	MsiI	MspI	MunI	MwoI
NciI	NcoI	NdeI	NheI	NlaIII	NlaIV	NsiI	NspI
NspBII	PstI	PvuII	RleAI	RsaI	SapI	Sau96I	Sau3AI
ScaI	ScrFI	SfaNI	SfcI	SpeI	StyI	TaqI	TaqII
TfiI	Tsp45I	Tsp509I	XcmI				

Enzymes that do not cut:

AatII	AccI	AflII	AgeI	Alw44I	ApaI	AscI	AvaI
AvaII	AvrII	BamHI	BanII	BcgI	BcgI	BglII	BsaAI
BsaHI	BscGI	BsiEI	BsiWI	BsmFI	Bsp1407I	BspEI	BspMI
BsrFI	BssHII	Bst1107I	BstXI	Bsu36I	DraI	DraIII	DrdI
EagI	Eam1105I	EciI	Eco4711I	Eco105I	EcoNI	Eco0109I	EcoRV
Esp3I	FseI	FspI	GdiI	HgaI	HgiEII	HincII	HindIII
HpaI	MluI	MlyI	MmeI	NoeI	NarI	NotI	NruI
NspV	PacI	Pfi1108I	PfIMI	PleI	PmeI	PmlI	PshAI
Psp5II	Psp1406I	PvuI	RcaI	RsrII	SacI	SacII	SalI
SexAI	SfiI	SgrAI	SmaI	SphI	SrfI	Sse8387I	SspI
StuI	Swal	ThaI	Tth111I	Tth1111I	VspI	XbaI	XhoI
XmnI							

FIG.4H

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/07295

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :A61K 39/00; C12N 15/12; C12P 21/02

US CL :424/185.1; 435/69.3, 320.1; 530/350

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 424/185.1; 435/69.3, 320.1; 530/350

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Medline, Biosis, Derwent, APS, GeneSeq, Swiss Prot, Pir 44, Embl-new6

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO, A, 93/25233 (UNIVERSITY OF CONNECTICUT) 23 December 1993, see entire document.	1, 4, 6-19, 21 and 23

☐

Further documents are listed in the continuation of Box C.

☐

See patent family annex.

* Special categories of cited documents:

A document defining the general state of the art which is not considered to be part of particular relevance

E earlier document published on or after the international filing date

L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

O document referring to an oral disclosure, use, exhibition or other means

P document published prior to the international filing date but later than the priority date claimed

T

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

X

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

Y

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

Z

document member of the same patent family

Date of the actual completion of the international search

10 SEPTEMBER 1995

Date of mailing of the international search report

11 OCT 1995

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

JULIE KRSEK-STAPLES

Telephone No. (703) 308-0196